STOCHASTIC PROCESSES AND COSMIC RADIATION

ON THE THEORY OF

STOCHASTIC PROCESSES

AND THEIR APPLICATION

TO THE THEORY OF

COSMIC RADIATION

BY

NIELS ARLEY

Assistant Professor of Physics
Institute of Theoretical Physics
University of Copenhagen

JOHN WILEY & SONS, INC., NEW YORK

PREFACE.

The purpose of the present paper is to investigate in further detail the so-called fluctuation problem in the theory of cosmic radiation. In the course of the calculations it turned out, however, that the theory of the stochastic processes required for this theory had not been treated before. The purely mathematical theory had, therefore, to be generalized. The author's interest in the mathematical problems was, furthermore, stimulated partly by the writing of a textbook on the theory of probability and its applications to statistics and the theory of errors, and partly by the appearance of a paper by the Swedish statistician fil. d:r. O. LUNDBERG on the theory of stochastic processes of exactly the type met with in the theory of cosmic radiation. As a result the paper falls into two separate parts, one purely mathematical and the other dealing with physics. Although the general results obtained in part I are utilized in part II, either part may, however, be read separately, each part having been written to form an entity.

I first of all wish to thank the director of the institute, Prof. dr. phil. NIELS BOHR for the great interest he has always shown my work and for the excellent working conditions he has given me at the institute.

Next, I wish to thank Prof. fil. d:r. HARALD CRAMÉR and his coworkers, especially fil. d:r. OVE LUNDBERG for many stimulating discussions on problems of probability. The theory of stochastic processes has been especially promoted by the Stockholm school under Prof. CRAMÉR and the mathematical part of the present paper represents a further generalization of some of the results obtained by Prof. CRAMÉR and his students.

My faithful coworker through many years, cand. mag. Mrs. BODIL TORNEHAVE, née ERIKSEN, I wish to thank most heartily for her invaluable assistance in the long and laborious numerical calculations which in the theory of cosmic rays, as in most modern physics, form

the necessary link between theory and experiment. I also wish to thank stud. mag. Mrs. INGER MØLLER who has assisted in some of the numerical calculations in chap. 4 and has drawn the figures. Mr. P. PROM, Sworn translator I thank for his kindness in controlling my English language and reading the proofs. For proof-reading I also thank mag. scient. K. RANDER BUCH and cand. mag. STIG BÜLOW. Furthermore, I wish to thank the printers, Messrs. Fr. Bagges kgl. Hofbogtrykkeri for their excellent work. Finally, I wish to thank my wife for her constant encouragement and great help in reading the proofs.

Institute of Theoretical Physics of the University of Copenhagen.

NIELS ARLEY

Copenhagen, March 1943.

CONTENTS.

LIST OF TABLES.

INTRODUCTION.

Throughout the history of mathematics one feature is conspicuous, namely *the intimate mutual interaction between the development of mathematics itself and its applications*, especially in physics. Sometimes mathematics has been a horse's head in advance having the theories fully developed before their application was required. We may mention, e. g., the RIEMANN geometry which later on became necessary for the mathematical formulation of general relativity theory and the theory of the HILBERT space which was subsequently applied in quantum theory. But likewise we meet with examples of mathematics being more or less behind the development of e. g. physics. The most striking example is offered by the creation of the infinitesimal calculus by NEWTON. Here even the necessary notions were lacking and NEWTON had to build for himself a quite new branch of mathematics from the ground. Another example is offered by the vector and tensor notions. It also happens that the mathematical interest in a problem is much furthered by physical and other applications. This is, e. g., the case of the FOURIER series and the problem of 'eigenvalues' and 'eigenfunctions'.

We are now in the same situation as regards the theory of *stochastic* or *random processes* which is dealt with in the first part of the present paper. These notions were first applied in the *mathematical theory of insurance risk*, but later on a purely mathematical theory was built up. This theory was, however, only developed up to a certain point and the applications then required the theory to be extended beyond this point, covering quite new processes. In actuarial science one type of extension was demanded, in physics another.

What is now meant by the term stochastic or random process? Before giving an answer to this question we shall remind the reader of the fact that whenever we are going to treat some real problem mathematically, whether in physics, in biology or in social sciences, we must necessarily start by simplifying the problem, having

recourse to some sort of a model representing those features of reality considered most essential for the problem in question. An important group of problems in the most varied fields of science may be reduced to the consideration of systems changing with time. For instance, in physics a diffusion process in which the position of some particles or the concentration of some material substance is changing with time, or the disintegration of atomic nuclei in which the state of the system changes with time. In astronomy the movement of the planets around the sun. In biology the growth of some organism or the development of a population of individuals. In actuarial science the changing number of claims on an insurance company for disbursements. In telephone theory the varying density of the traffic. In economics the varying demands and prices of the market and so on. A multitude of other examples might be mentioned.

In some of the examples quoted above the appropriate model for the mathematical description is a *well-determined model* as e. g. in classical mechanics where the state of a system is uniquely determined for all times if the state is known at one definite time. In other of the examples it is, however, necessary to apply a *statistical description*, either because it is *practically* impossible to give a *causal* description of the system as is e. g. the case in statistical mechanics, or because it is *fundamentally* impossible as in quantum theory. In both cases we have to use *statistical models* in the mathematical description. In such models the knowledge of the system being in a certain state at one definite time does *not* determine uniquely the state of the system for all times, but only a certain *probability* for each of the possible states. Furthermore, it is, as a rule, characteristic of the statistical models that the time parameter can move only in the *positive* time direction, whereas in a well-determined model the state is determined both for all subsequent and for all previous times. We may here stress the fact that *the question of the possibility of applying a well-determined or a statistical model in the mathematical description of some real process is quite independent of whether the process itself is of a causal or of a random nature.*

In earlier times only such statistical models were applied in which the system considered might change in discrete moments, $t_1, t_2, \ldots$ only. Sometimes a limit operation was afterwards carried out, whereby an asymptotic form of the distribution was obtained. In this way the Gauss'ian distribution was e. g. deduced for certain

diffusion problems as an asymptotic expression for the binomial distribution. In more recent times we have, however, been led to consider statistical models in which this limit operation is carried out already in the definition proper of the model and the time parameter may thus vary *continuously*. Such statistical models are called *stochastic* or *random processes*. We thereby attain the great mathematical advantage that the probabilities of the model become governed by ordinary or partial differential equations with respect to the time parameter. Furthermore, the expressions which in the previous theory had only asymptotic validity now constitute the exact solutions themselves.

It has turned out that the new statistical models with a continuously varying time parameter have a wide field of application, being most valuable in the mathematical treatment of statistical problems in the most varied branches of science. Among these applications two types of stochastic processes are conspicuous. Considering the examples quoted above of systems changing with time and being of a statistical nature we see that sometimes the stochastic variables entering into the model change *continuously* as is e. g. the case in the diffusion problems or in the biological problems of growth of an organism. Sometimes they may, however, change only by *finite amounts at discrete moments* as is e. g. the case in the problem of the changing number of claims on an insurance company for disbursements or in the biological problem of the changing numbers of individuals present in a population. The latter problem exhibits great resemblance to the problems of primary interest to us, namely the development of *cosmic ray showers* by the passage of very energetic cosmic ray particles through matter. Here the thickness of the layer penetrated corresponds to the time in the biological analogy. Furthermore, the *electrons* and light quanta, the *photons*, constituting the shower, may disappear, 'die', or give birth to new generations of electrons and photons, 'propagate', just as is the case with the corresponding individuals in the biological population.

Stochastic processes in the sense just outlined were first considered by BACHELIER[1]) and LUNDBERG[2]). They have been especially studied by various authors for the purpose of the mathematical theory of insurance risk. A complete and mathematically rigorous theory was, however, first given in the fundamental paper of

[1]) BACHELIER (1900) and (1912).

[2]) LUNDBERG, F. (1909).

KOLMOGOROFF[1]). Later on, FELLER[2]) presented the theory in a general form, which contains continuous and discontinuous processes as special cases. For the purpose of the insurance theory LUNDBERG[3]) has generalized FELLER's theory of discontinuous processes. Finally, *we aim in the first part of the present paper (chapters 1–3) at giving a further generalization of Feller's and Lundberg's theories for the purpose of the application of stochastic processes to the theory of cosmic ray showers.*

The mathematical theory of part I forms the theoretical background for the applications of the stochastic processes we are going to make in part II (chapters 4–6). In this part we discuss in detail the *fluctuation problem in the theory of the cosmic ray showers* already touched upon above. These showers, being the most conspicuous phenomena in cosmic ray physics, consist as mentioned of very energetic electrons and photons which give rise to the creation of new generations of electrons and photons by the passage through matter. It is the purpose of the theory to work out in detail the mean numbers of particles in the showers and the fluctuations about these mean numbers, thus obtaining estimates of the probabilities of the showers containing different numbers of particles. These probabilities can be directly compared with experimental curves, the *Rossi curves*.

For several reasons it is desirable to work out such a detailed theory of the shower phenomenon. In the first place, it is important to ascertain whether the behaviour of electrons and photons may be entirely accounted for by the present relativistic quantum theory, or whether this theory breaks down at high energies so that quite new ideas have to be introduced. In the second place, it is of decisive importance to have as detailed knowledge as possible of this electron-photon part, the *soft component*, of the cosmic rays in order to obtain a clear picture of the properties which must be ascribed to the other part, the *hard component*, consisting of much more penetrating particles, the *mesons*. The nature of these new particles, hitherto not found in laboratory experiments, is now the main problem in cosmic ray physics. For instance, it is extremely important for the theory to know whether *neutral* mesons exist in the cosmic rays.

1) KOLMOGOROFF (1931).
2) FELLER (1937).
3) LUNDBERG, O. (1940).

The theory of the shower phenomenon was first given simultaneously in the fundamental papers of BHABHA and HEITLER[1]) and CARLSON and OPPENHEIMER[2]) and has later on formed the subject of several other investigations. The theory has been enlarged in two papers by the author[3]), specially dealing with the theory of *coincidence experiments* yielding the Rossi curves mentioned above. *It is the purpose of the second part of the present paper to examine more closely the theoretical calculation of these Rossi curves by investigating the rôle played by the fluctuation problem.*

Finally, we shall discuss the experimental material available for a *quantitative* comparison between theory and experiment. By comparisons between theoretical and experimental curves we are usually surprised when smaller or greater differences are found between the two sets of curves. Bearing in mind the long series of more or less rough approximations underlying the theoretical curves on the one hand and the interpretation of the experimental curves on the other, we think that we ought in fact to look at things the other way round and rather be surprised at the wonderful fact that the two sets of curves even agree *qualitatively*, giving results of the same order of magnitude. Especially in cosmic ray physics we are, however, reminded of the ancient fact that nature is really much more simple than might be expected, the various errors tending to compensate rather than to magnify each other. *It is, namely, a fundamental, empirical law of nature that the region of validity of the simplifying approximations, of which any theory must necessarily make use, is always far wider than might be justified by theoretical arguments.*

---

1) BHABHA and HEITLER (1937).

2) CARLSON and OPPENHEIMER (1937).

3) ARLEY (1938).
ARLEY and ERIKSEN (1940).

PART I

MATHEMATICAL THEORY

CHAPTER 1.

Survey of the Theory of Stochastic Processes.

§ **1.1.** *By a* **stochastic process** *we understand a process in which enter one or more stochastic variables*[1]) *the probability distribution of which depends on one continuously varying parameter.* The stochastic variables we shall denote by x_1, x_2, and so on. The parameter we shall denote by t because it will be convenient to interpret it as denoting *time*. Each of the stochastic variables $x_1(t)$, $x_2(t)$, . . . entering into the process is assumed to move in a Euclidean space of one or more dimensions. If only one stochastic variable enters we shall speak of a **one-dimensional**, otherwise, of a **multi-dimensional stochastic process,** of dimension m.

We shall now in this chapter discuss the general stochastic processes according to the theories of KOLMOGOROFF[2]) and FELLER[3]). In the practical applications of stochastic processes two types of processes are met with, as mentioned in the introduction, the *continuous* and the *discontinuous* processes. Theoretically we may also have a mixture of these types, but such *mixed* processes have so far not been met with. The *continuous* processes we meet with e. g. in physics when considering *diffusion problems*. Here the position, as a function of time, of a colloidal particle suspended in a liquid is the stochastic variable. The characteristic property of such processes consists in small stochastic changes taking place incessantly. The *discontinuous* processes we meet with e. g. in biology when considering a population consisting of one or more species of individuals each of which can be created, propagate or die. Here the numbers of individuals of the various species present at a given time are the stochastic variables. The characteristic property of such processes consists in the stochastic changes occurring at discrete

---

[1]) Also called *random* or *statistical* variables.

[2]) KOLMOGOROFF (1931).

[3]) FELLER (1937).

moments and with finite values. We therefore say in this case that the stochastic variables *jump* from one state to another.

As mentioned in the introduction we must always when describing some real happening have recourse to the construction of simplified models of reality, introducing some kind of a model. If this model is *well-determined*, as e. g. in classical physics, it will nearly always be the case that the state of the system considered will be determined for all times t if only the state is known at one definite time s. And if, furthermore, some knowledge is also required of the antecedent before the time s of the system this knowledge may as a rule be taken into account by the specification of the values at the time s of some further quantities, parameters, describing the antecedent, as e. g. the velocities in classical physics. In the same way it will nearly always be the case when the model employed is a *stochastic* model, that the probability distribution of the system considered will be determined for all times t ($> s$) if the state is known at one definite time s[1]). For such processes KOLMOGOROFF has introduced the term **stochastically definite.** Also in this case it may often be possible to take the antecedent into account by introducing parameters in a proper way. *We shall in the following always tacitly assume the stochastic process under consideration to be a stochastic definite.* We now first outline the theory of one-dimensional processes.

§ 1.2. The general theory of stochastic processes is based on the modern axiomatic development of the theory of probability as first put in systematic form by KOLMOGOROFF[2]). An event is here labelled by a point set A in a Euclidean space E with a finite number of dimensions, and all sets considered are assumed to be Borel sets. As the probability of the event labelled A we may simply define an arbitrary real set function $P(A)$, the **probability function,** satisfying the well-known axioms

---

[1]) We may, however, even in simple practical problems meet with systems in which the knowledge of the antecedent enters directly (cf. § 4.5).

[2]) KOLMOGOROFF (1933). Cf. also CRAMÉR (1937). Regarding the definitions and theorems of the elementary theory of probability we may refer to any textbook on probability, e. g. that of CRAMÉR just mentioned. In the following we shall, however, refer Danish readers to the textbook of ARLEY and BUCH (1943) quoted as A & B.

$$0 \leq P(A) \leq 1 , \tag{1}$$

$$P(E) = 1 \tag{2}$$

and if no two of A_1, A_2, ... have common points

$$P(A_1+A_2+\ldots) = P(A_1)+P(A_2)+\ldots \qquad (A_i \cdot A_k = O) . \tag{3}$$

If for some value t our stochastic variable $\boldsymbol{x}(t)$ assumes a value lying inside a certain region A of the given Euclidean space, we say that the event $\boldsymbol{x}(t) \subset A$ has occurred. Assuming now our stochastic process to be *stochastically definite*, the conditioned probability of an event $\boldsymbol{x}(t) \subset A$, relative to the hypothesis that the variable assumes the value x at the moment s, $\boldsymbol{x}(s) = x$, is uniquely determined for every moment $t > s$ by a **conditioned or relative probability function** of the type

$$P(A, t;\ x, s) = \text{Probability } \big(\boldsymbol{x}(t) \subset A;\ \boldsymbol{x}(s) = x\big) . \tag{4}$$

If the variable $\boldsymbol{x}(t)$ has, furthermore, a given initial value, we may also speak of the *absolute* probability of the event $\boldsymbol{x}(t) \subset A$ and thus of an **absolute probability function**

$$P(A, t) = \text{Probability } \big(\boldsymbol{x}(t) \subset A\big) . \tag{5}$$

If this probability function is defined for $t = s$ and if the process is stochastically definite, then it follows from the general composition rules for probabilities that

$$P(A, t) = \int_E P(A, t;\ \xi, s)\, d_X P(X, s) . \tag{6}$$

Here the integral is a *Lebesgue-Stieltjes integral* taken over the whole space E. We note that the differential here and in the following refers to the set X, a set being in the following denoted by capital letters, and that ξ denotes a point corresponding to X.

From the definition of the set function $P(A, t;\ x, s)$, defined in (4) and characterizing the stochastic process in question, and from the general composition rules for probabilities it follows that P for all $t \geq s$ must satisfy the following conditions

$$0 \leq P(A, t;\ x, s) \leq 1 \qquad \text{for all } t \geq s , \tag{7}$$

$$\int_E d_A P(A, t;\ x, s) = 1 \qquad \text{for all } t \geq s \tag{8}$$

and

$$P(A, t;\ x, s) = \int_E P(A, t;\ \xi, \tau) d_X P(X, \tau;\ x, s) \qquad (9)$$

for all $t \geq s$
„ „ τ in $s \leq \tau \leq t$.

The relation (9) is called the **Chapman-Kolmogoroff equation.** For $t = s$ P is reduced to the **unit-distribution** which, using a generalization of the well-known δ-symbol of KRONECKER, introduced by the author in another connection[1]), may be expressed by

$$\Delta(x \subset A) = \begin{cases} 1 & \text{for } x \subset A \\ 0 & \text{„ } x \subset E - A . \end{cases} \qquad (10)$$

Now we make the natural assumption that P is a *continuous function* of both s and t ($\geq s$). We thus have

$$\lim_{t \to s} P(A, t;\ x, s) = \lim_{t \to s} P(A, t;\ x, s) = \Delta(x \subset A) . \qquad (11)$$

Furthermore, P is assumed to be *B-measurable with respect to x.*

On the other hand, every set function satisfying the **fundamental conditions** (7)-(9) and (11) defines for $t \geq s$ a stochastically definite process. We thus see, that the problem of defining analytically such a process is equivalent to the problem of finding solutions of the CHAPMAN-KOLMOGOROFF equation (9) which are continuous in s and t, B-measurable in x and satisfy the fundamental conditions (7), (8) and (11). We shall not here discuss this problem in general but only in case the process is of the *discontinuous type*. For the discussion of continuous (and mixed) processes we refer to KOLMOGOROFF[2]), FELLER[3]) and the review of the theory given by KHINTCHINE[4]).

For the analytical description of discontinuous stochastic processes it is convenient to introduce two new functions, the **intensity function** $p(x, t)$ and the **relative transition probability function** $\Pi(A;\ x, t)$. Here $p(x, t)\Delta t$ is an asymptotic expression of the probability of a stochastic change of the variable taking place in the interval between t and $t + \Delta t$ when the variable assumes the value x at the moment t. Next, $\Pi(A;\ x, t)$ is the conditioned probability of

[1]) Cf. ARLEY (1938), reviewed in chap. 5 inf.
[2]) KOLMOGOROFF (1931).
[3]) FELLER (1937).
[4]) KHINTCHINE (1933).

the variable assuming a value in the set A at the time $t+\Delta t$, $x(t+\Delta t) \subset A$, relative to the hypothesis that a stochastic change of the variable from the value x has taken place during the interval between t and $t+\Delta t$. From these definitions it follows that p and Π satisfy the following conditions

$$p(x, t) \geqq 0 \qquad \text{for all } x,\ t \geqq s, \tag{12}$$

$$0 \leqq \Pi(A;\ x, t) \leqq 1 \qquad \text{for all } A \text{ and } x,\ t \geqq s, \tag{13}$$

$$\Pi(A;\ x, t) \equiv 0 \qquad \text{for all } x \subset A,\ t \geqq s, \tag{14}$$

and

$$\int_E d_X \Pi(X;\ x, t) = 1 \qquad \text{for all } x,\ t \geqq s. \tag{15}$$

Furthermore, we must for small values of Δt have the asymptotic expression

$$P(A, t+\Delta t;\ x, t) = \\ (1-p(x, t)\Delta t)\Delta(x \subset A)+\Pi(A;\ x, t)p(x, t)\Delta t+o(\Delta t)\ . \tag{16}$$

(16) together with (7)–(9) and (11) constitute the *fundamental conditions for a discontinuous, stochastically definite process.* FELLER[1]) adds the further conditions that the *p and Π functions are continuous in t and B-measurable in x* and that the intensity function is *uniformly bounded* with respect to x in every finite time interval

$$p(x, t) \leqq K(t) \qquad \text{for all } x,\ t \geqq s. \tag{17}$$

This last condition is, however, too restrictive for most practical applications of the theory. LUNDBERG[2]) has generalized FELLER's theory to embrace the processes being of importance in actuarial science, but this generalization is not sufficiently wide to cover also the processes needed in the theory of cosmic radiation. It is our purpose in chap. 2 and 3 to give such a generalization.

Replacing t by $t+\Delta i$ and τ by t in the CHAPMAN-KOLMOGOROFF equation (9) and inserting (16) for the first factor in the integral we obtain

---

[1]) FELLER (1937).

[2]) LUNDBERG (1940).

$$P(A, t+\Delta t;\ x, s) = \int_E P(A, t+\Delta t;\ \xi, t) d_X P(X, t;\ x, s) =$$

$$\int_E (1-p(\xi, t)\Delta t)\Delta(\xi \subset A) d_X P(X, t;\ x, s) +$$

$$\Delta t \int_E \Pi(A;\ \xi, t) p(\xi, t) d_X P(X, t;\ x, s) + o(\Delta t) =$$

$$P(A, t;\ x, s) - \Delta t \int_A p(\xi, t) d_X P(X, t;\ x, s) +$$

$$\Delta t \int_E \Pi(A;\ \xi, t) p(\xi, t) d_X P(X, t;\ x, s) + o(\Delta t)\ . \qquad (18)$$

From (18) it follows that the first partial derivative of P with respect to t exists, because

$$\lim_{\Delta t \to 0} \frac{P(A, t+\Delta t;\ x, s) - P(A, t;\ x, s)}{\Delta t} = \frac{\partial}{\partial t} P(A, t;\ x, s) =$$

$$-\int_A p(\xi, t) d_X P(X, t;\ x, s) + \int_E \Pi(A;\ \xi, t) p(\xi, t) d_X P(X, t;\ x, s)\ . \quad (19)$$

In the same way we obtain, inserting (16) for the second factor in (9) with $\tau = s+\Delta s$, that the first partial derivative of P with respect to s exists and is given by an equation analogous with (19)

$$\frac{\partial}{\partial s} P(A, t;\ x, s) =$$

$$p(x, s) P(A, t;\ x, s) - p(x, s) \int_E P(A, t;\ \xi, s) d_X \Pi(X;\ x, s)\ . \quad (20)$$

These equations we shall call the **fundamental equations** of the process.

Feller[1]) *has now proved the following main theorem. Feller's theorem: Under the conditions* (12)–(15) *and* (17) *of the functions* $p(x, t)$ *and* $\Pi(A;\ x, t)$ *and assuming them to be continuous in t and B-measurable in x there exists one and the same unique solution of the fundamental equations* (19) *and* (20). *This solution is continuous in s and t, B-measurable in x and satisfies the five fundamental conditions* (7)–(9), (11) *and* (16) *of a discontinuous process. The solution thus constitutes a conditioned probability function which characterizes a stochastically definite process of the discontinuous type.*

[1]) Feller (1937).

§ 1.3. Even in the general form in which we have presented the theory in the preceding paragraph FELLER's theorem can be extended to more general processes for which the essential condition (1.2.17) is no longer fulfilled[1]). In the discontinuous processes met with in the practical applications of the theory it will, however, *most frequently be the case that the stochastic variable* $\boldsymbol{x}(t)$ *can assume only an enumerable manifold of values.* Whether or not these values are integers we can then always denote the possible states of the variable by the integers $n = 0, 1, 2, \ldots$. Such stochastic processes are called **Markoff chains.** The *set* function $P(A, t; x, s)$ characterizing a Markoff chain is thus reduced to a *point* function $P(n, t;\ n', s)$ in which n' denotes the *initial* and n the *final* state. In the same way the set function $\Pi(A\,; x, t)$ is reduced to the point function $\Pi(n\,; n', t)$. The Lebesgue-Stieltjes integrals of § 1.2 are now reduced to ordinary, finite or infinite, sums and the fundamental equations (1.2.19) and (1.2.20) thus read

$$\frac{\partial}{\partial t} P(n, t;\ n', s) =$$

$$-p(n, t)P(n, t;\ n', s) + \sum_{n''=0}^{\infty}{}' \Pi(n; n'', t)p(n'', t)P(n'', t; n', s) \qquad (1)$$

and

$$\frac{\partial}{\partial s} P(n, t;\ n', s) =$$

$$P(n, t;\ n', s)p(n', s) - \sum_{n''=0}^{\infty}{}' P(n, t;\ n'', s)\Pi(n'';\ n', s)p(n', s)\ . \qquad (2)$$

Furthermore, the fundamental conditions (1.2.7)–(1.2.9), (1.2.11) and (1.2.16) read

$$\lim_{t \to s} P(n, t; n', s) = \lim_{t \leftarrow s} P(n, t; n', s) = \delta_{n,n'} = \begin{cases} 1 \text{ for } n = n' \\ 0 \quad \text{,,} \quad n \neq n' \end{cases} \qquad (\mathrm{I})$$

$$P(n, t + \Delta t;\ n', t) =$$
$$(1 - p(n', t)\Delta t)\delta_{n,n'} + \Pi(n\,; n', t)p(n', t)\Delta t + o(\Delta t) \qquad (\mathrm{II})$$

$$P(n, t;\ n', s) = \sum_{n''=0}^{\infty}{}' P(n, t;\ n'', \tau)\ P(n'', \tau;\ n', s) \qquad (\mathrm{III})$$

$$0 \leq P(n, t;\ n', s) \leq 1 \qquad (\mathrm{IV})$$

---

[1]) The author intends to describe such a generalization at another place.

$$\sum_{n=0}^{\infty} P(n, t;\ n', s) = 1\ . \tag{V}$$

For a *finite* number of possible states KOLMOGOROFF[1]) has shown that if the p and Π functions are assumed to be continuous in t, then a unique solution exists, being the same for (1) and (2) and satisfying the fundamental conditions (I)–(V). We note that KOLMOGOROFF's theorem is obviously a special case of FELLER's theorem stated above. For a discussion of the finite case we may, furthermore, refer to FRÉCHET[2]). In the next chapter we shall give a detailed discussion of the fundamental equations (1) and (2) and the fundamental conditions (I)–(V).

We observe that in all quantum mechanical perturbation problems involving a consideration of time, stochastically definite processes are met with, which are, however, described in a way other than that discussed above[3]). The reason is that in quantum theory we do not consider the probabilities themselves but complex quantities satisfying equations analogous to (1) and (2) and the numerical squares of which are interpreted as probabilities. Furthermore, these quantities are governed by a kind of generating function, the wave function, being the expansion coefficients of this function in a series of 'eigenfunctions' of the unperturbed system in question. From the theory of the wave equation it follows that the probabilities in question are uniquely determined and satisfy all claims which must be exacted from them in order that they shall represent a stochastic process, viz. conditions (I)–(V). For the stochastic processes used in quantum theory no special theory is thus required.

§ 1.4. In the preceding paragraphs we have reviewed the theory of one-dimensional stochastic processes, i.e. processes in which only one stochastic variable enters. We cannot, however, be content with such processes because in the practical applications of the theory we shall need processes in which *two* stochastic variables enter (cf. chap. 4). Now the theory is, of course, formally the same whether two or more variables enter and we shall, therefore, in chap. 3 generalize the one-dimensional theory to a theory of multi-dimensional, discontinuous stochastic processes. Biological examples of such processes we meet with when e. g. considering a population consisting of individuals belonging to *different* species.

[1]) KOLMOGOROFF (1931). Cf. § 2.8 inf.

[2]) FRÉCHET (1938).

[3]) Cf. e. g. HEITLER (1936) chap. III § 9.3.

Each individual may, we assume, disappear, be created or propagate and by these processes one or more individuals of the same or of other species may at the same time disappear or be created. Our stochastic variables then show the numbers of individuals of the different species present at the time t.

Let the number of the stochastic variables entering into the process be m. An m-dimensional stochastically definite process is then uniquely described by a conditioned probability function of the type[1])

$$P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) =$$
$$\text{Probability}\left(\boldsymbol{x}_1(t) \subset A_1, \ldots, \boldsymbol{x}_m(t) \subset A_m;\ \boldsymbol{x}_1(s) = x_1, \ldots, \boldsymbol{x}(s) = x_m\right) \quad (1)$$

denoting the probability of the simultaneous occurrence of the m events $\boldsymbol{x}_1(t)$ assuming a value in the set A_1, and so on, relative to the hypothesis that the m events $\boldsymbol{x}_1(s) = x_1$, and so on, have occurred simultaneously.

In analogy with (1.2.7)–(1.2.9) and (1.2.11) this m-dimensional set function, assumed to be continuous in t and s and B-measurable in $x_1, \ldots, x_m$, must satisfy the relations

$$0 \leq P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) \leq 1 \qquad \text{for all } t \geq s\,, \quad (2)$$

$$\int_{E_1} \cdots \int_{E_m} d_A P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) = 1 \qquad \text{for all } t \geq s\,, \quad (3)$$

in which the integral is an *m-dimensional Lebesgue-Stieltjes integral,* the **Chapman-Kolmogoroff equation**

$$P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) =$$
$$\int_{E_1} \cdots \int_{E_m} P(A_1, \ldots, A_m, t;\ \xi_1, \ldots, \xi_m, \tau)\, d_X P(X_1, \ldots, X_m, \tau;\ x_1, \ldots, x_m, s) \quad (4)$$

for all $t \geq s$
„ „ τ in $s \leq \tau \leq t$

and finally

$$\lim_{t \to s} P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) =$$

$$\lim_{t \leftarrow s} P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) = \Delta(x_1 \subset A_1) \cdots \Delta(x_m \subset A_m)\,. \quad (5)$$

[1]) We remind the reader of the fact that (cf. § 1.2) each of the variables $x_1, \ldots, x_m$ denote a point in a Euclidean space of a finite dimension, not necessarily the same for all the variables.

In analogy with the intensity and the relative transition probability functions we introduce the **m-dimensional intensity function** $p(x_1, \ldots, x_m, t)$ and the **m-dimensional relative transition probability function** $\Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t)$. Here $p(x_1, \ldots, x_m, t)\Delta t$, assumed to be continuous in t and B-measurable in $x_1, \ldots, x_m$, is an asymptotic expression of the probability of a stochastic change of the variables taking place in the interval between t and $t+\Delta t$ when the variables assume the values $x_1, \ldots, x_m$, respectively, at the time t. Next $\Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t)$, assumed to be continuous in t and B-measurable in $x_1, \ldots, x_m$, is the conditioned probability of the variables assuming values in the sets $A_1, \ldots, A_m$, respectively, at the time $t+\Delta t$, relative to the hypothesis that a stochastic change of the variables from the values $x_1, \ldots, x_m$ has taken place during the interval between t and $t+\Delta t$. In analogy with (1.2.12)–(1.2.15) the p and Π functions must satisfy the relations

$$p(x_1, \ldots, x_m, t) \geq 0 \qquad \text{for all } x_1, \ldots, x_m \quad \text{,, ,, } t \geq s\,, \tag{6}$$

$$0 \leq \Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t) \leq 1 \qquad \text{for all } A_1, \ldots, A_m \text{ and } x_1, \ldots, x_m \quad \text{,, ,, } t \geq s \tag{7}$$

$$\Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t) = 0 \quad \text{for all } t \geq s\,, \tag{8}$$

when $x_1 \subset A_1, \ldots, x_m \subset A_m$ are all satisfied, and

$$\int_{E_1} \cdots \int_{E_m} d_A \Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t) = 1 \quad \text{for all } x_1, \ldots, x_m \quad \text{,, ,, } t \geq s\,. \tag{9}$$

In analogy with (1.2.16) we must for small values of Δt have the asymptotic expression

$$\begin{aligned} &P(A_1, \ldots, A_m, t+\Delta t; x_1, \ldots, x_m, t) = \\ &(1-p(x_1, \ldots, x_m, t)\Delta t)\, I(x_1 \subset A_1) \cdots I(x_m \subset A_m) + \\ &\Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t)\, p(x_1, \ldots, x_m, t)\Delta t + o(\Delta t)\,. \end{aligned} \tag{10}$$

The relations (2)–(5) and (10) constitute the **fundamental conditions** *for an m-dimensional, discontinuous, stochastically definite process*. Inserting (10) for the first factor in the integral in the CHAPMAN-KOLMOGOROFF equation (4) we obtain, in the same way as in § 1.2, that the first partial derivatives of P with respect to t and s both exist and satisfy the following equations which are analogous with (1.2.19) and (1.2.20), respectively,

$$\frac{\partial}{\partial t} P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s) =$$

$$-\int_{A_1} \cdots \int_{A_m} p(\xi_1, \ldots, \xi_m, t) d_X P(X_1, \ldots, X_m, t;\ x_1, \ldots, x_m, s) +$$

$$\int_{E_1} \cdots \int_{E_m} \Pi(A_1, \ldots, A_m;\ \xi_1, \ldots, \xi_m, t) p(\xi_1, \ldots, \xi_m, t) \cdot$$

$$d_X P(X_1, \ldots, X_m, t;\ x_1, \ldots, x_m, s) \quad (11)$$

and

$$\frac{\partial}{\partial s} P(A_1, \ldots, A_m, t; x_1, \ldots, x_m, s) =$$

$$p(x_1, \ldots, x_m, s) P(A_1, \ldots, A_m, t; x_1, \ldots, x_m, s) -$$

$$p(x_1, \ldots, x_m, s) \int_{E_1} \cdots \int_{E_m} P(A_1, \ldots, A_m, t;\ \xi_1, \ldots, \xi_m, s) \cdot$$

$$d_X \Pi(X_1, \ldots, X_m;\ x_1, \ldots, x_m, s) . \quad (12)$$

These equations we shall call the **fundamental equations** of the process.

As was the case with the one-dimensional theory, the m-dimensional theory is seldom applied to practical problems in this general form. It will most frequently be the case that each of the stochastic variables $x_1(t), \ldots, x_m(t)$ can assume only an enumerable manifold of values. Whether or not these values are integers, we can always denote the possible states of the variables by a set of integers, $n_1, \ldots, n_m = 0, 1, 2, \ldots$. The m-dimensional *set* function $P(A_1, \ldots, A_m, t;\ x_1, \ldots, x_m, s)$ characterizing such **Markoff chains** is reduced to a *point* function $P(n_1, \ldots, n_m, t; n_1', \ldots, n_m', s)$ in which the m-dimensional point $(n_1', \ldots, n_m') = (n')$ denotes the *initial* and $(n_1, \ldots, n_m) = (n)$ the *final* state. In the same way the set function $\Pi(A_1, \ldots, A_m; x_1, \ldots, x_m, t)$ is reduced to the point function $\Pi(n_1, \ldots, n_m; n_1', \ldots, n_m', t)$. The m-dimensional Lebesgue-Stieltjes integrals are now reduced to m-dimensional sums and the fundamental equations (11) and (12) thus read

$$\frac{\partial}{\partial t} P((n), t;\ (n'), s) = -p((n), t) P((n), t;\ (n'), s) +$$

$$\sum_{n_1''=0}^{\infty} \cdots \sum_{n_m''=0}^{\infty} \Pi((n);\ (n''), t) p((n''), t) P((n''), t;\ (n'), s) \quad (13)$$

and

$$\frac{\partial}{\partial s}P\big((n), t;\ (n'), s\big) = P\big((n), t;\ (n'), s\big)p\big((n'), s\big) -$$

$$\sum_{n_1''=0}^{\infty} \cdots \sum_{n_m''=0}^{\infty} P\big((n), t;\ (n''), s\big)\Pi\big((n'');\ (n'), s\big)p\big((n'), s\big) . \qquad (14)$$

The fundamental conditions (2)–(5) and (10) finally read

$$\lim_{t\to s} P\big((n), t;\ (n'), s\big) = \lim_{t\leftarrow s} P\big((n), t;\ (n'), s\big) = \delta_{n_1, n_1'} \cdots \delta_{n_m, n_m'} \qquad (I)$$

$$P\big((n), t+\Delta t;\ (n'), t\big) =$$

$$\big(1 - p\big((n'), t\big)\Delta t\big)\delta_{n_1, n_1'} \cdots \delta_{n_m, n_m'} + \Pi\big((n);\ (n'), t\big)p\big((n'), t\big)\Delta t + o(\Delta t) \quad (II)$$

$$P\big((n), t;\ (n'), s\big) =$$

$$\sum_{n_1''=0}^{\infty} \cdots \sum_{n_m''=0}^{\infty} P\big((n), t;\ (n''), \tau\big)P\big((n''), \tau;\ (n'), s\big) \qquad (III)$$

$$0 \leq P\big((n), t;\ (n'), s\big) \leq 1 \qquad (IV)$$

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} P\big((n), t;\ (n'), s\big) = 1 . \qquad (V)$$

In chap. 3 we shall discuss the fundamental equations (13) and (14) and the fundamental conditions (I)–(V).

CHAPTER 2.

The general Theory of one-dimensional Stochastic Processes.

§ 2.1. We shall now discuss the theory of general stochastic processes of the discontinuous type in which the stochastic variable can assume only an enumerable manifold of values. As we have seen in chap. 1 such processes are described by two infinite systems of simultaneous linear differential equations of the first order

$$\frac{\partial}{\partial t} P(n, t;\ n', s) =$$

$$-p(n, t)P(n, t;\ n', s) + \sum_{n''=0}^{\infty} \Pi(n;\ n'', t)p(n'', t)P(n'', t;\ n', s) \quad (1)$$

$$(n, n' = 0, 1, 2, \ldots\ ;\ t \geqq s)$$

and

$$\frac{\partial}{\partial s} P(n, t;\ n', s) =$$

$$P(n, t;\ n', s)p(n', s) - \sum_{n''=0}^{\infty} P(n, t;\ n'', s)\Pi(n'';\ n', s)p(n', s) \quad (2)$$

$$(n, n' = 0, 1, 2, \ldots\ ;\ t \geqq s),$$

in which we shall call (2) the equation **adjointed** to (1).

The symbols entering into (1) and (2) have the following meaning:

$P(n, t;\ n', s)$ *is the conditioned, or relative, probability of the stochastic variable assuming the value n at the time t relative to the hypothesis that it assumes the value n' at the time s.*

$p(n, t)\Delta t$ *is an asymptotic expression of the probability of a stochastic change of the variable taking place in the interval between t and $t+\Delta t$ when the variable assumes the value n at the time t; $p(n, t)$ is called the* **intensity function.**

$\Pi(n; n', t)$ *is the conditioned probability of the variable assuming the value n at the time* $t+\Delta t$ *relative to the hypothesis that a stochastic change of the variable from the state* n' *has taken place during the interval between t and* $t+\Delta t$. *We shall call* $\Pi(n; n', t)$ *the* **relative transition probability from** *the state* n' **to** *the state n.*

By means of these definitions the equations (1) and (2) have an immediate 'anschaulich' interpretation, representing 'continuity equations' in the positive and negative time direction, respectively, for the 'probability mass', expressing that this 'mass' can neither be created nor disappear. On the left side of (1) we have the rate of change with respect to t and this must equal the 'probability current' which 'flows into' the n'th state from all the other states, i. e. second term on right, minus the 'probability current' which 'flows out' from the n th state to all the other states, i. e. first term on right. In the same way the left side of (2) represents the rate of change with respect to s and this must equal the 'probability current' which 'flows out' from the state n' to all the other states, i. e. first term on right, minus the 'probability current' which 'flows into' the state n' from all the other states, i. e. second term on right.

From the definition of the p and Π functions it follows that

$$p(n, t) \geq 0 \quad \text{for all } n = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geq s, \tag{3}$$

$$0 \leq \Pi(n; n', t) \leq 1 \quad \text{for all } n, n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geq s, \tag{4}$$

$$\Pi(n'; n', t) = 0 \quad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geq s \tag{5}$$

and

$$\sum_{n=0}^{\infty} \Pi(n; n', t) = 1 \quad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geq s. \tag{6}$$

It is now the object of our theory to show, as discussed in chap. 1, *that if all the functions* $p(n, t)$ *and* $\Pi(n; n', t)$ *are given, satisfy* (3)–(6) *and are assumed to be* c o n t i n u o u s *for all* $t \geq s$, *then* (1) *and* (2) *each has one and only one solution which is the same for both systems and which satisfies the following conditions being necessary and sufficient for the solutions denoting a probability distribution for a discontinuous stochastic process*

$$\lim_{t \to s} P(n, t; n', s) = \lim_{t \leftarrow s} P(n, t; n', s) = \delta_{n,n'} = \begin{cases} 1 & \text{for } n = n' \\ 0 & \text{,, } n \neq n' \end{cases} \quad \text{(I)}$$

for all $n, n' = 0, 1, 2, \ldots$

$$P(n, t+\Delta t; n', t) = \begin{cases} 1 - p(n', t)\Delta t + o(\Delta t) & \text{for } n = n' \\ \Pi(n; n', t)p(n', t)\Delta t + o(\Delta t) & \text{for } n \neq n' \end{cases} \quad \text{(II)}$$

for all $t \geqq s$

$$P(n, t; n', s) = \sum_{n''=0}^{\infty} P(n, t; n'', \tau) P(n'', \tau; n', s) \quad \text{(III)}$$

for all $n, n' = 0, 1, 2, \ldots$
,, ,, τ in $s \leqq \tau \leqq t$

(CHAPMAN-KOLMOGOROFF's equation)

$$0 \leqq P(n, t; n', s) \leqq 1 \quad \text{for all } n, n' = 0, 1, 2, \ldots \text{ ,, ,, } t \geqq s \quad \text{(IV)}$$

$$\sum_{n=0}^{\infty} P(n, t; n', s) = 1 \quad \text{for all } n' = 0, 1, 2, \ldots \text{ ,, ,, } t \geqq s . \quad \text{(V)}$$

Before proceeding we shall write our equations in a more compact form using the *matrix* symbolism[1]). We first introduce the **distribution matrix**

$$\boldsymbol{P}(t, s) = \{(n|\boldsymbol{P}(t, s)|n')\} = \{P(n, t; n', s)\} = \begin{Bmatrix} P(0, t; 0, s), & P(0, t; 1, s), & \ldots \\ P(1, t; 0, s), & P(1, t; 1, s), & \ldots \\ \vdots & \vdots & \end{Bmatrix}. \quad (7)$$

Next we introduce the **intensity matrix** as the diagonal matrix

$$\boldsymbol{p}(t) = \{(n|\boldsymbol{p}(t)|n')\} = \begin{Bmatrix} p(0, t), & 0 \;, & 0 \;, & \ldots \\ 0 \;, & p(1, t), & 0 \;, & \ldots \\ 0 \;, & 0 \;, & p(2, t), & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{Bmatrix} \quad (8)$$

---

[1]) Cf. § 7.1. We here use the matrix symbolism introduced by DIRAC since it is especially convenient and symmetric. Matrices themselves will be denoted by clarendon letters, $\boldsymbol{P}, \boldsymbol{p}, \boldsymbol{\Pi}$ etc. and their elements by $(n|\boldsymbol{P}|n')$, $(n|\boldsymbol{p}|n')$, $(n|\boldsymbol{\Pi}|n')$ etc. n is called the **row index** and n' the **column index.** Furthermore, we shall write $\boldsymbol{P} = \{(n|\boldsymbol{P}|n')\}$ etc.

and finally the **relative transition matrix**

$$\boldsymbol{\Pi}(t) = \{(n|\boldsymbol{\Pi}(t)|n')\} = \begin{Bmatrix} \Pi(0;0,t), & \Pi(0;1,t), & \ldots \\ \Pi(1;0,t), & \Pi(1;1,t), & \ldots \\ \vdots & \vdots & \end{Bmatrix}. \tag{9}$$

Furthermore, we introduce a matrix operation of which we shall frequently make use. We define

$$\boldsymbol{\Sigma} \cdot \boldsymbol{M} = \left\{ \sum_{n=0}^{\infty} (n|\boldsymbol{M}|n') \right\} = \{(\boldsymbol{\Sigma} \cdot \boldsymbol{M}|n')\}$$

and (10)

$$\boldsymbol{M} \cdot \boldsymbol{\Sigma} = \left\{ \sum_{n'=0}^{\infty} (n|\boldsymbol{M}|n') \right\} = \{(n|\boldsymbol{M} \cdot \boldsymbol{\Sigma})\}.$$

These matrices with *one* row and *infinite* columns, respectively *infinite* rows and *one* column, we shall call the **column sum** and **row sum matrix,** respectively[1]).

We can now write our **fundamental equations** (1) and (2) in the form

$$\frac{\partial}{\partial t} \boldsymbol{P}(t,s) = -\boldsymbol{p}(t) \cdot \boldsymbol{P}(t,s) + \boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t) \cdot \boldsymbol{P}(t,s) = \boldsymbol{A}(t) \cdot \boldsymbol{P}(t,s) \tag{1a}$$

and

$$\frac{\partial}{\partial s} \boldsymbol{P}(t,s) = \boldsymbol{P}(t,s) \cdot \boldsymbol{p}(s) - \boldsymbol{P}(t,s) \cdot \boldsymbol{\Pi}(s) \cdot \boldsymbol{p}(s) = -\boldsymbol{P}(t,s) \cdot \boldsymbol{A}(s) \tag{2a}$$

or

$$\frac{\partial}{\partial s} \boldsymbol{P}^*(t,s) = -\boldsymbol{A}^*(s) \cdot \boldsymbol{P}^*(t,s)\,. \tag{2b}$$

(2b) shows that *equation* (2) *is of exactly the same form as* (1).

The **operator matrix** introduced in (1a) and (2a)

$$\boldsymbol{A}(t) = -\boldsymbol{p}(t) + \boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t) = (\boldsymbol{\Pi}(t) - \boldsymbol{1}) \cdot \boldsymbol{p}(t) \tag{11}$$

has the elements

[1]) We note that $\boldsymbol{\Sigma} \cdot \boldsymbol{M}$ can also be interpreted as the matrix product between the infinite matrix $\boldsymbol{\Sigma} = \{1, 1, 1, \ldots\}$ and $\boldsymbol{M}$. In the same way $\boldsymbol{M} \cdot \boldsymbol{\Sigma}$ can be interpreted as the matrix product between $\boldsymbol{M}$ and $\boldsymbol{\Sigma}^*$. We thus have e. g. that $\boldsymbol{\Sigma} \cdot (|\boldsymbol{K}| \cdot |\boldsymbol{M}|) = (\boldsymbol{\Sigma} \cdot |\boldsymbol{K}|) \cdot |\boldsymbol{M}|$ and $(|\boldsymbol{K}| \cdot |\boldsymbol{M}|) \cdot \boldsymbol{\Sigma} = |\boldsymbol{K}| \cdot (|\boldsymbol{M}| \cdot \boldsymbol{\Sigma})$ when $\boldsymbol{K}$ and $\boldsymbol{M}$ are arbitrary matrices for which the products exist.

$$(n|\boldsymbol{A}(t)|n') = \begin{cases} \Pi(n;\, n', t)p(n', t) \geq 0 & \text{for} \quad n \neq n' \\ -p(n', t) \leq 0 & \text{for} \quad n = n' \end{cases} \tag{12}$$

and satisfies, due to (6),

$$\boldsymbol{\Sigma} \cdot \boldsymbol{A}(t) = \left\{ \sum_{n=0}^{\infty} (n|\boldsymbol{A}(t)|n') \right\} = \boldsymbol{0} \quad \text{for all } n' = 0, 1, 2, \ldots \tag{13}$$

The **fundamental conditions** (I)–(V) can now be written

$$\lim_{t \to s} \boldsymbol{P}(t, s) = \lim_{t \leftarrow s} \boldsymbol{P}(t, s) = \boldsymbol{1} \tag{Ia}$$

$$\boldsymbol{P}(t + \Delta t, t) = \boldsymbol{1} + \boldsymbol{A}(t)\Delta t + \boldsymbol{o}(\Delta t) \qquad \text{for all } t \geq s \tag{IIa}$$

$$\boldsymbol{P}(t, s) = \boldsymbol{P}(t, \tau) \cdot \boldsymbol{P}(\tau, s) \qquad \text{for all } s \leq \tau \leq t \tag{IIIa}$$

$$\boldsymbol{0} \leq \boldsymbol{P}(t, s) \leq \{1\} \qquad \text{for all } t \geq s \tag{IVa}$$

$$\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s) = \{1\} \qquad \text{for all } t \geq s\,. \tag{Va}$$

As mentioned in chap. 1 FELLER[1]) has proved that (I)–(V) are satisfied on the assumption that the intensity function is uniformly bounded, i. e.

$$p(n, t) \leq K(t) \qquad \text{for all } n = 0, 1, 2, \ldots \tag{14}$$

Next LUNDBERG[2]) has investigated under what conditions for $p(n, t)$ (I)–(V) remain satisfied on the assumption that the process is **elementary** which means that the variable can change only by $+1$ in each change, i. e.

$$(n|\boldsymbol{\Pi}(t)|n') = f(n', t)\delta_{n, n'+1}\,. \tag{15}$$

We shall now extend these theories, though without attempting to give the most general theory applicable for all probability operators. We shall in the following make the essential assumption that the probability operator, besides being continuous, is **absolutely exponentiable in some interval** (s, t), which means (cf. chap. 7) that the matrix function

$$\exp[\boldsymbol{K}(t - s)] = \sum_{\nu=0}^{\infty} \boldsymbol{K}^{\nu} \frac{(t - s)^{\nu}}{\nu!} \tag{16}$$

exists, when

$$\boldsymbol{K} = \max_{s \leq \tau \leq t} |\boldsymbol{A}(\tau)|\,. \tag{17}$$

[1]) FELLER (1937). Cf. § 2.9 inf.

[2]) LUNDBERG (1940). Cf. § 2.9 inf.

The stochastic processes met with in practice will—presumably—always fulfil this condition. In chap. 7 we have shown that this is so, at any rate, for four important classes of probability operators which embrace as special cases those of KOLMOGOROFF, FELLER and LUNDBERG (cf. especially the summary in § 7.6).

First, we shall show, in the following paragraphs, that equations of the type (1a), i. e. both (1) and (2) as shown by (2b), have for an *arbitrary*, continuous operator matrix $\boldsymbol{A}(t)$ satisfying (16) one and only one solution which always automatically will satisfy the adjointed equation and fulfil the conditions (Ia)–(IIIa). When $\boldsymbol{A}(t)$ is, furthermore, a probability operator, i. e. of the type (11), the solution will also fulfil the condition (IVa). Finally, we shall discuss the condition (Va) which is the only essential condition, the fulfilment of which will impose certain restrictions on $\boldsymbol{A}(t)$.

§ 2.2. First of all we show that if our equation (2.1.1a) actually *has* any solution, it can have only *one*. This theorem has, of course, nothing to do with probabilities and we therefore prove it generally.

Let

$$\frac{d}{dx}\boldsymbol{Y}(x) = \boldsymbol{A}(x)\cdot\boldsymbol{Y}(x) \tag{1}$$

be an arbitrary, real or complex, infinite system of simultaneous linear differential equations of the first order[1]) with the initial condition

$$\boldsymbol{Y}(x_0) = \boldsymbol{C} \tag{2}$$

in which $\boldsymbol{C}$ is an arbitrary constant matrix. We only assume the, real or complex, operator matrix $\boldsymbol{A}(x)$ to be continuous and absolutely exponentiable in an interval (x_0, x), i. e. to satisfy (2.1.16). From (1) it follows that $\boldsymbol{Y}$ is continuous and assuming, furthermore, the product $\boldsymbol{A}\cdot\boldsymbol{Y}$ to be *uniformly* convergent in x, it then follows that also $\boldsymbol{Y}'$ is continuous. We shall prove that (1) can at most have *one* solution $\boldsymbol{Y}(x)$. If, namely, (1) had two different solutions, $\boldsymbol{Y}_1(x)$ and $\boldsymbol{Y}_2(x)$, with the same initial condition (2), then

$$\boldsymbol{Y}(x) = \boldsymbol{Y}_1(x)-\boldsymbol{Y}_2(x) \tag{3}$$

would also be a solution, because

[1]) For simplicity's sake we assume in the following the variable x to be real, but the theorems and proofs can be immediately generalized to the case in which x is a complex variable on a regular curve without double points in the complex plane.

$$\frac{d}{dx}\boldsymbol{Y} = \frac{d}{dx}\boldsymbol{Y}_1 - \frac{d}{dx}\boldsymbol{Y}_2 = \boldsymbol{A}\cdot\boldsymbol{Y}_1 - \boldsymbol{A}\cdot\boldsymbol{Y}_2 = \boldsymbol{A}\cdot\boldsymbol{Y}. \tag{4}$$

Due to (2) $\boldsymbol{Y}(x)$ would, furthermore, have the initial condition

$$\boldsymbol{Y}(x_0) = \boldsymbol{Y}_1(x_0) - \boldsymbol{Y}_2(x_0) = \boldsymbol{C} - \boldsymbol{C} = \boldsymbol{0}. \tag{5}$$

We have, consequently, only to show that if a solution of (1) is zero at a point x_0, then it must remain identically zero through the whole region of definition of $\boldsymbol{A}(x)$. This we prove by the classical method of *iteration*.

Let $\boldsymbol{K}$ be given by (2.1.17) and let

$$\boldsymbol{G} = \max_{x_0 \leq t \leq x} |\boldsymbol{Y}(t)|. \tag{6}$$

We then have from (1) that

$$|\boldsymbol{Y}'| \leq \boldsymbol{K}\cdot|\boldsymbol{Y}| \leq \boldsymbol{K}\cdot\boldsymbol{G} \qquad \text{for all } t \text{ in } x_0 \leq t \leq x. \tag{7}$$

$\boldsymbol{Y}'$ being continuous, we may integrate (7) and, using (5), we obtain

$$|\boldsymbol{Y}| \leq \boldsymbol{K}\cdot\boldsymbol{G}(x-x_0) \qquad \text{for all } t \text{ in } x_0 \leq t \leq x. \tag{8}$$

Inserting (8) into (7) gives

$$|\boldsymbol{Y}'| \leq \boldsymbol{K}^2\cdot\boldsymbol{G}(x-x_0) \tag{9}$$

which integrated gives

$$|\boldsymbol{Y}| \leq \boldsymbol{K}^2\cdot\boldsymbol{G}\frac{(x-x_0)^2}{2!}, \tag{10}$$

and so on ad infinitum. We thus have that

$$|\boldsymbol{Y}| \leq \boldsymbol{K}^\nu \frac{(x-x_0)^\nu}{\nu!}\cdot\boldsymbol{G} \qquad \begin{array}{l}\text{for all } \nu = 1, 2, 3, \ldots \\ \text{,, ,, } t \text{ in } x_0 \leq t \leq x.\end{array} \tag{11}$$

Now the right hand side of (11) tends to zero when ν tends to infinity as we have assumed $\boldsymbol{A}$ to be absolutely exponentiable in the interval (x_0, x). (11) can, consequently, be satisfied for all values of ν only if

$$\boldsymbol{Y} \equiv \boldsymbol{0} \quad \text{in } x_0 \leq t \leq x, \tag{12}$$

q. e. d.

§ 2.3. We next prove the theorem of existence of the equation (2.1.1a). This theorem has no more to do with probabilities than the theorem on uniqueness just proved and we shall, therefore, again

prove it generally. Let again

$$\frac{d}{dx} \boldsymbol{Y} = \boldsymbol{A} \cdot \boldsymbol{Y} \tag{1}$$

be an arbitrary, real or complex, infinite system of simultaneous linear differential equations of the first order[1]) with the initial condition, $\boldsymbol{C}$ being an arbitrary constant matrix,

$$\boldsymbol{Y}(x_0) = \boldsymbol{C}\,. \tag{2}$$

We again assume the, real or complex, operator matrix $\boldsymbol{A}(x)$ to be continuous and absolutely exponentiable in the interval (x_0, x). Next we assume the product

$$\mathbf{exp}[\boldsymbol{K}(x - x_0)] \cdot |\boldsymbol{C}| \tag{3}$$

to exist, in which $\boldsymbol{K}$ is given by (2.1.17). We note that (3) being a *power* series in $(x - x_0)$, the product $\boldsymbol{K}^\nu \cdot \mathbf{exp}[\boldsymbol{K}(x - x_0)] \cdot |\boldsymbol{C}|$, being the ν'th differential coefficient of $\mathbf{exp}[\boldsymbol{K}(x - x_0)] \cdot |\boldsymbol{C}|$, exists for all values of ν. We now prove, again using the classical method of *iteration*, that (1) has always a solution. This proof runs, of course, in exactly the same way as the proof of uniqueness in § 2.2.

Before starting we shall write (1) in another form. All quantities being continuous (cf. an observation in § 2.2) we can integrate both sides of (1) from x_0 to x and, using (2), (1) is thus identical with

$$\boldsymbol{Y}(x) = \boldsymbol{C} + \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{Y}(t)\,dt\,. \tag{4}$$

We now put

$$\boldsymbol{Y}(x) = \sum_{\nu=0}^{\infty} \boldsymbol{Y}_\nu(x) \tag{5}$$

in which

$$\begin{aligned} \boldsymbol{Y}_0(x) &= \boldsymbol{C} \\ \text{and} \qquad \boldsymbol{Y}_\nu(x) &= \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{Y}_{\nu-1}(t)\,dt \qquad \nu = 1, 2, 3, \ldots\,, \end{aligned} \tag{6}$$

and shall prove (a) that the series (5) exists, (b) that it is a solution of (4), i. e. of (1).

Firstly, all products in (6) exist and are continuous, as all the defining series are *uniformly* convergent due to (3). Consequently, all the functions $\boldsymbol{Y}_\nu$ exist, are continuous and, furthermore, differentiable with continuous differential coefficients. Next it follows, using (2.1.17), that

[1]) We assume again x to be a real variable cf. [1]) p. 36.

$$|Y_1| \leqq \int_{x_0}^{x} K \cdot |C| dt = K \cdot |C|(x-x_0)\,,$$
$$|Y_2| \leqq \int_{x_0}^{x} K \cdot K \cdot |C|(t-x_0)dt = K^2 \cdot |C| \frac{(x-x_0)^2}{2!}\,, \tag{7}$$

and so on ad infinitum. We thus have

$$|Y_\nu(t)| \leqq K^\nu \cdot |C| \frac{(x-x_0)^\nu}{\nu!} \qquad \text{for all } \nu = 0, 1, 2, \ldots \quad \text{,, ,, } t \text{ in } x_0 \leqq t \leqq x\,. \tag{8}$$

Inserting (8) into (5) we have, as all summations are absolutely convergent and their order therefore irrelevant,

$$|Y(t)| \leqq \sum_{\nu=0}^{\infty} |Y_\nu| \leqq \sum_{\nu=0}^{\infty} K^\nu \cdot |C| \frac{(x-x_0)^\nu}{\nu!} = \exp[K(x-x_0)] \cdot |C| \tag{9}$$
$$\text{for all } t \text{ in } x_0 \leqq t \leqq x\,.$$

(9) shows that the series (5) is absolutely convergent the convergence being uniform in the interval $x_0 \leqq t \leqq x$. $Y(x)$, *consequently, exists and is continuous.*

Inserting (6) into (5) we next find

$$Y(x) = C + \sum_{\nu=1}^{\infty} \int_{x_0}^{x} A(t) \cdot Y_{\nu-1}(t) dt = C + \int_{x_0}^{x} \left(\sum_{\nu=0}^{\infty} A(t) \cdot Y_\nu(t) \right) dt =$$
$$C + \int_{x_0}^{x} \left(A(t) \cdot \sum_{\nu=0}^{\infty} Y_\nu(t) \right) dt = C + \int_{x_0}^{x} A(t) \cdot Y(t) dt\,. \tag{10}$$

This calculation is legitimate because, firstly, the convergence in ν is *uniform* with respect to t due to (2.1.17), (3) and (9) and, secondly, all summations are *absolutely* convergent. From (3) and (9) it follows, furthermore, that $A \cdot Y$ is *continuous* and (10) shows, consequently, that $Y(x)$, being an integral of a continuous function, is differentiable with the continuous differential coefficient

$$\frac{d}{dx} Y(x) = A(x) \cdot Y(x) \tag{11}$$

which is actually (1), q. e. d.

§ 2.4. By a **fundamental solution** of our equation $Y' = A \cdot Y$ we understand a solution having the property that *any* solution (cf. however below) is the product of this solution and a constant

matrix. We shall now show that if we put $C = 1$ in (2.3.6) we obtain a fundamental solution, which we shall denote by

$$\boldsymbol{F}(x, x_0) = \sum_{\nu=0}^{\infty} \boldsymbol{F}_\nu(x, x_0) \tag{1}$$

in which

$$\boldsymbol{F}_0(x, x_0) = \mathbf{1}$$

and (2)

$$\boldsymbol{F}_\nu(x, x_0) = \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{F}_{\nu-1}(t, x_0)\,dt \qquad \nu = 1, 2, 3, \ldots .$$

We first note some properties of $\boldsymbol{F}(x, x_0)$. As we have seen in § 2.3 this function is continuous and has a continuous differential coefficient with respect to x which satisfies

$$\frac{\partial}{\partial x}\boldsymbol{F}(x, x_0) = \boldsymbol{A}(x) \cdot \boldsymbol{F}(x, x_0) \quad \text{or} \quad \boldsymbol{F}(x, x_0) = \mathbf{1} + \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{F}(t, x_0)\,dt . \tag{3}$$

Furthermore, we have seen that the series (1) is absolutely and uniformly convergent with respect to x as (1) is majorized by the series (2.3.9), i. e.

$$|\boldsymbol{F}(t, x_0)| \leq \sum_{\nu=0}^{\infty} |\boldsymbol{F}_\nu(t, x_0)| \leq \sum_{\nu=0}^{\infty} \boldsymbol{K}^\nu \frac{(x - x_0)^\nu}{\nu!} = \exp[\boldsymbol{K}(x - x_0)] \tag{4}$$

for all t in $x_0 \leq t \leq x$.

From (3) it is seen that

$$\lim_{x_0 \to x} \boldsymbol{F}(x, x_0) = \lim_{x_0 \leftarrow x} \boldsymbol{F}(x, x_0) = \mathbf{1} . \tag{5}$$

Next we see that (2) can also be written in another way:

$$\boldsymbol{F}_\nu(x, x_0) = \int_{x_0}^{x} \boldsymbol{A}(t) \cdot \boldsymbol{F}_{\nu-1}(t, x_0)\,dt =$$

$$= \int_{x \geq t_\nu \geq} \cdots \int_{\geq t_1 \geq x_0} \boldsymbol{A}(t_\nu) \cdots \boldsymbol{A}(t_1)\,dt_\nu \ldots dt_1 = \int_{x_0}^{x} \boldsymbol{F}_{\nu-1}(x, t) \cdot \boldsymbol{A}(t)\,dt \tag{6}$$

for all $\nu = 1, 2, 3, \ldots$.

Introducing the right hand side of (6) into (1) we have, for the same reasons as in (2.3.10),

$$\boldsymbol{F}(x, x_0) = 1 + \sum_{\nu=1}^{\infty} \int_{x_0}^{x} \boldsymbol{F}_{\nu-1}(x, t) \cdot \boldsymbol{A}(t)\,dt =$$

$$1 + \int_{x_0}^{x} \Big(\sum_{\nu=0}^{\infty} \boldsymbol{F}_{\nu}(x, t)\Big) \cdot \boldsymbol{A}(t)\,dt = 1 + \int_{x_0}^{x} \boldsymbol{F}(x, t) \cdot \boldsymbol{A}(t)\,dt. \tag{7}$$

(7) shows that $\boldsymbol{F}(x, x_0)$ is also a continuous and differentiable function with respect to x_0 with the continuous differential coefficient

$$\frac{\partial}{\partial x_0} \boldsymbol{F}(x, x_0) = -\boldsymbol{F}(x, x_0) \cdot \boldsymbol{A}(x_0) \tag{8}$$

which is exactly the equation being adjointed to (3).

Finally, we note that $\boldsymbol{A}(t)$ being continuous, it follows from (1) and (2) that

$$\boldsymbol{F}(x + \Delta x, x) = 1 + \int_{x}^{x+\Delta x} \boldsymbol{A}(t)\,dt + \sum_{\nu=2}^{\infty} \boldsymbol{F}_{\nu}(x + \Delta x, x) =$$
$$1 + \boldsymbol{A}(x)\Delta x + \boldsymbol{o}(\Delta x). \tag{9}$$

It is now obvious that if $\boldsymbol{C}$ is an arbitrary matrix which is *independent* of x and for which the product

$$\exp[\boldsymbol{K}(x - x_0)] \cdot |\boldsymbol{C}| \tag{10}$$

exists[1]), then the product

$$\boldsymbol{Y}(x, x_0) = \boldsymbol{F}(x, x_0) \cdot \boldsymbol{C} \tag{11}$$

exists and is again a solution of our fundamental equation (3). All the summations in (11) being absolutely convergent we can, namely, multiply by $\boldsymbol{C}$ before we sum over ν and (11) is, thus, identical with the solution (2.3.5). We note that this $\boldsymbol{Y}$ does not, of course, satisfy the adjointed equation (8), as was the case for $\boldsymbol{F}$, because $\boldsymbol{C}$ may depend on x_0. Due to (11) and (5) we have, namely,

$$\boldsymbol{Y}(x_0, x_0) = \boldsymbol{F}(x_0, x_0) \cdot \boldsymbol{C} = \boldsymbol{C}. \tag{12}$$

On the other hand, if $\boldsymbol{Y}(x, x_0)$ is an arbitrary solution of (3) which has the property that, $\boldsymbol{G}$ being given by (2.2.6), the product

$$\exp[\boldsymbol{K}(x - x_0)] \cdot \boldsymbol{G} \tag{13}$$

exists, then $\boldsymbol{Y}$ must be given by (11) with $\boldsymbol{C}$ determined by (12). In

[1]) It is evident that $\boldsymbol{C}$ cannot in general be a quite arbitrary infinite matrix as $\boldsymbol{F}(x, x_0)$ is in general *infinite* and the product (11) need not, consequently, exist at all.

fact, $F(x, x_0) \cdot Y(x_0, x_0)$ exists due to (13) and is, as we have just seen, another solution of (3) which for $x = x_0$ coincides with Y. As shown in § 2.2 they must, consequently, coincide for all values of x. *Our solution* (1) *has, thus, in fact the property of being a fundamental solution.*

We stress that in the following we shall always by a solution understand a solution in the limited sense that (13) *is fulfilled*[1]).

§ 2.5. We have now seen in the preceding paragraphs that our fundamental equation

$$\frac{d}{dx} Y = A \cdot Y \tag{1}$$

always has one and only one solution under the assumption that A is continuous and absolutely exponentiable in the interval (x_0, x). Furthermore, we have seen that our fundamental solution $F(x, x_0)$ given in (2.4.1), i. e. the solution with

$$F(x_0, x_0) = 1 , \tag{2}$$

will automatically also be the—unique—solution of the adjointed equation (2.4.8) and, moreover, satisfy (2.4.5) and (2.4.9). We see that the two last relations are identical with (2.1.Ia) and (2.1.IIa), respectively. Before showing that our fundamental solution will also automatically satisfy (2.1.IIIa) it will be convenient to show that instead of being expressed by the series (2.4.1), it can also be expressed in the form of a *product-integral.*

The notion of the product-integral, which was introduced for *finite* systems of differential equations by VOLTERRA[2]) and discussed by SCHLESINGER[3]), RASCH[4]) and others, is arrived at in the following way.

Let us consider an interval $x_0 \leq \tau \leq t$ ($\leq x$) and let

$$x_0 < x_1 < x_2 < \ldots < x_{m-1} < x_m = t \leq x, \quad \Delta_i = x_{i+1} - x_i \tag{3}$$

[1]) We suppose that no solutions exist other than those satisfying (13), but we have not investigated this point more closely as it is of no interest in our applications of the theory.

[2]) VOLTERRA (1887), (1888) and (1902).

[3]) SCHLESINGER (1908) and (1922).

[4]) RASCH (1930) chap. II. The theory is in this paper represented by means of the matrix symbolism. In the present paragraph we shall generalize RASCH's representation from finite to infinite equations.

be a division of the interval. This particular division we shall denote by 'm'. If the division is very fine, i. e. $\max\{\Delta_i\}$ $i = 0, 1, 2, \ldots, m-1$ is very small, which fact we shall express by saying that 'm' is very large, then we shall have

$$F(x_{i+1}, x_0) = F(x_i, x_0) + \Delta_i F'(x_i, x_0) + o(\Delta_i) . \tag{4}$$

Neglecting the rest term $o(\Delta_i)$[1]) and inserting (1) into (4) we have approximately

$$F(x_{i+1}, x_0) = F(x_i, x_0) + \Delta_i A(x_i) \cdot F(x_i, x_0) = \big(1 + A(x_i)\Delta_i\big) \cdot F(x_i, x_0) . \tag{5}$$

Repeating this m times we have approximately[2]), using (2),

$$F(t, x_0) = \big(1 + A(x_{m-1})\Delta_{m-1}\big) \cdots \big(1 + A(x_1)\Delta_1\big) \cdot \big(1 + A(x_0)\Delta_0\big) = \prod_{i=0}^{m-1} \big(1 + A(x_i)\Delta_i\big) \tag{6}$$

which product exists for all values of m due to our assumption of A being absolutely exponentiable in (x_0, x). We thus expect that making the division finer and finer, i. e. making $\max\{\Delta_i\} \to 0$ which fact we shall simply express by 'm' $\to \infty$, the operator in (6) shall converge towards a limit equal to our fundamental solution (2.4.1). This limit is called the **product-integral** and we shall denote it by the symbol[3])

$$\lim_{'m' \to \infty} \prod_{i=0}^{m-1} \big(1 + A(x_i)\Delta_i\big) = \mathcal{P}_{x_0}^{t} \big(1 + A(\tau)d\tau\big) = F(t, x_0) . \tag{7}$$

In order to prove the convergence in (7) we first write the operator in (6) in a way first indicated by Peano[4]) in the case of A being

[1]) This means assuming Δ_i so small that each of the F-functions can be approximated throughout Δ_i by its tangent in x_i.

[2]) We note that all products here and in the following have to be read *from right to left*. If we had defined A by $Y' = Y \cdot A$ all products should have been read from left to right.

[3]) Volterra writes it $\int_{x_0}^{t} A(\tau)d\tau$, Schlesinger $\int_{x_0}^{t} (A(\tau)d\tau + 1)$ and Rasch $\prod_{x_0}^{t} (1 + A(\tau)d\tau)$. Just as $\int$ is a symbol which shall remind of Σ but be different from it, the symbol of the product-integral is to be reminiscent of both $\int$ and Π but be different from both. We therefore consider a stylized P the most appropriate symbol.

[4]) Peano (1888).

finite. Performing all the multiplications in (6) and using (a) that $\mathbf{1}$ commutes with all matrices and (b) that the associative law is valid for the products occurring in (6), when $\boldsymbol{A}$ is absolutely exponentiable, we obtain

$$\prod_{i=0}^{m-1}(\mathbf{1}+\boldsymbol{A}(x_i)\Delta_i) = \tag{8}$$

$$(\mathbf{1}+\boldsymbol{A}(x_{m-1})\Delta_{m-1})\cdots(\mathbf{1}+\boldsymbol{A}(x_1)\Delta_1)\cdot(\mathbf{1}+\boldsymbol{A}(x_0)\Delta_0) = \sum_{\nu=0}^{m}\boldsymbol{F}_\nu^{(m)}(t, x_0)$$

in which

and
$$\boldsymbol{F}_0^{(m)}(t, x_0) = \mathbf{1} \tag{9}$$

$$\boldsymbol{F}_\nu^{(m)}(t, x_0) = \sum_{m-1 \geq i_\nu > \ldots > i_1 \geq 0} \ldots \sum \boldsymbol{A}(x_{i_\nu})\cdots\boldsymbol{A}(x_{i_1})\Delta_{i_\nu}\cdots\Delta_{i_1} \qquad \nu = 1, 2, 3, \ldots, m.$$

As $\boldsymbol{A}$ is continuous, $\boldsymbol{F}_\nu^{(m)}(t, x_0)$ is for each value of $\nu \leq m$ a continuous function of the ν variables $x_{i_1}, \ldots, x_{i_\nu}$. If we now make the limit 'm' $\to \infty$, then it follows from the definition of a ν-multiple integral of a continuous function that the limit exists and is given by

$$\lim_{'m' \to \infty}\boldsymbol{F}_\nu^{(m)}(t, x_0) = \int_{t > t_\nu > \ldots > t_1 > x_0}\ldots\int \boldsymbol{A}(t_\nu)\cdots\boldsymbol{A}(t_1)\,dt_\nu\ldots dt_1 = \boldsymbol{F}_\nu(t, x_0) \tag{10}$$

for all $\nu = 1, 2, 3, \ldots$.

Comparing with (2.4.6) we see that the limit in (10) is exactly the function $\boldsymbol{F}_\nu(t, x_0)$ introduced in (2.4.2). From (9) it follows, using (2.1.17) and the polynomial formula,[1])

$$\left|\boldsymbol{F}_\nu^{(m)}(t, x_0)\right| \leq K^\nu \sum_{m-1 \geq i_\nu > \ldots > i_1 \geq 0}\ldots\sum \Delta_{i_\nu}\cdots\Delta_{i_1} \leq$$

$$K^\nu\frac{1}{\nu!}\sum_{\alpha_0+\ldots+\alpha_{m-1}=\nu}\ldots\sum\frac{\nu!}{\alpha_0!\cdots\alpha_{m-1}!}\Delta_0^{\alpha_0}\cdots\Delta_{m-1}^{\alpha_{m-1}} =$$

$$K^\nu\frac{1}{\nu!}(\Delta_0+\ldots+\Delta_{m-1})^\nu \leq K^\nu\frac{(x-x_0)^\nu}{\nu!} \tag{11}$$

for all 'm'
,, ,, t in $x_0 \leq t \leq x$
,, ,, $\nu \leq m = 1, 2, 3, \ldots$.

[1]) We observe that due to the fact that $\int_{t > t_\nu \geq \ldots > t_1 \geq x_0}\ldots\int dt_\nu\ldots dt_1 = \frac{(t-x_0)^\nu}{\nu!}$ it is seen that the majorization in (11) cannot be sharpened if it shall be valid for all values of m.

It is now obvious that the limit in (7) exists and that it is equal to $F(t, x_0)$ given in (2.4.1) because we have, using (2.4.4) and (11),

$$(p|\sum_{\nu=0}^{\infty} F_\nu(t, x_0) - \prod_{i=0}^{m-1}(1+A(x_i)\Delta_i)|q) \leq$$

$$(p|\sum_{\nu=0}^{n} F_\nu(t, x_0) - \sum_{\nu=0}^{n} F_\nu^{(m)}(t, x_0)||q) + (p|\sum_{\nu=n+1}^{\infty} |F_\nu(t, x_0)||q) +$$

$$(p|\sum^{m} |F_\nu^{(m)}(t, x_0)||q) \leq$$

$$(p|\sum_{\nu=1}^{n} |F_\nu(t, x_0) - F_\nu^{(m)}(t, x_0)||q) + 2(p|\sum_{\nu=n+1}^{\infty} K^\nu \frac{(x-x_0)^\nu}{\nu!}|q) \tag{12}$$

for each fixed $p, q = 0, 1, 2, \ldots$
,, all 'm'
,, ,, t in $x_0 \leq t \leq x$
,, ,, $n < m = 1, 2, 3, \ldots$.

For an arbitrary $\varepsilon > 0$ we can now, for each fixed value of p and q, first choose n so large that the second sum on the right side of (12) is smaller than $\frac{\varepsilon}{3}$ for all 'm' for which $m > n$. Next we can choose m ($> n$) so large that each of the n terms in the first sum on the right side of (12) is less than $\frac{\varepsilon}{3n}$. We have thus proved the convergence of the product-integral and that its limit is in fact given by (7). We note that (12) shows that the convergence process is uniform in each interval $x_0 \leq t \leq x$.

From the definition of the product-integral it now follows immediately that whenever $x_0 \leq x_1 \leq x$, where (x_0, x) is an interval in which A is absolutely exponentiable,

$$\mathcal{P}_{x_1}^{x}(1+A(t)dt) \cdot \mathcal{P}_{x_0}^{x_1}(1+A(t)dt) = \mathcal{P}_{x_0}^{x}(1+A(t)dt)$$

i. e. (13)

$$F(x, x_0) = F(x, x_1) \cdot F(x_1, x_0) \ .$$

Due to (2), or directly from the definition of the product-integral, it follows that this formula is also valid for $x = x_0$. This shows that the matrix $\mathcal{P}_{x_0}^{x}(1+A(t)dt)$ has a reciprocal given by

$$\left(\prod_{x_0}^{x}(1+A(t)dt)\right)^{-1}=\prod_{x}^{x_0}(1+A(t)dt). \qquad (14)$$

We can now recapitulate the results obtained so far. We have seen that equations of the type (1) *have always one and only one solution in the interval* (x_0, x), *for a given initial condition, if the operator matrix* $A(x)$ *is continuous and absolutely exponentiable in* (x_0, x). *The solution corresponding to the initial condition* (2) *has, in particular, always the properties expressed in* (2.4.5), (2.4.8), (2.4.9) *and* (13). *This shows that the fundamental equation* (2.1.Ia) *has always one and only one solution which will automatically also be the–unique–solution of the adjointed equation* (2.1.2a) *and automatically satisfy the fundamental conditions* (2.1.Ia)–(2.1.IIIa). *These conditions are, consequently, not limited to the theory of probability, but will always be satisfied. We thus see that the product-integral is the ideal mathematical tool for the theory of discontinuous processes!*

§ 2.6. We shall now give a few important properties of the product-integral of which we shall make use later on. We first observe that if $A(x)$ and $F_1(x, x_0) = \int_{x_0}^{x} A(t)dt$ *commute*, i. e.

$$A \cdot F_1 = F_1 \cdot A , \qquad (1)$$

then the product-integral can be evaluated by simple quadrature. If, namely, a matrix function $F(x)$ commutes with its differential coefficient $F'(x)$, then we have

$$\frac{d}{dx}(F(x))^n = F' \cdot F \cdots F + F \cdot F' \cdot F \cdots F + \ldots + F \cdots F \cdot F' = n(F(x))^{n-1} \cdot F'(x) . \qquad (2)$$

As $A(x) = \frac{\partial}{\partial x} F_1(x, x_0)$, (2.4.2) in this case consequently reads

$$F_2(x, x_0) = \int_{x_0}^{x}\left(\frac{\partial}{\partial t} F_1(t, x_0)\right) \cdot F_1(t, x_0)dt = \frac{1}{2!}(F_1(x, x_0))^2 , \qquad (3)$$

and so on, and thus generally

$$F_\nu(x, x_0) = \int_{x_0}^{x}\left(\frac{\partial}{\partial t} F_1(t, x_0)\right) \cdot \frac{1}{(\nu-1)!}(F_1(t, x_0))^{\nu-1}dt = \frac{1}{\nu!}(F_1(x, x_0))^\nu$$

$$\text{for all } \nu = 0, 1, 2, \ldots . \qquad (4)$$

Inserting (4) into the series (2.4.1) for the product-integral we find

$$\mathop{\mathcal{P}}_{x_0}^{x}(1+A(t)dt) = F(x, x_0) = \sum_{\nu=0}^{\infty}\frac{1}{\nu!}(F_1(x, x_0))^{\nu} =$$

$$\exp[F_1(x, x_0)] = \exp\left[\int_{x_0}^{x} A(t)dt\right] , \tag{5}$$

the exponential existing due to the assumption of A being absolutely exponentiable in the interval (x_0, x). An important special case is offered when $A(x)$ is the product of a *constant* matrix A and a *scalar* function $a(x)$. We then have

$$F_1(x, x_0) = A\int_{x_0}^{x} a(t)dt \tag{6}$$

which matrix commutes, of course, with A. (5) is then simply reduced to

$$\mathop{\mathcal{P}}_{x_0}^{x}(1+A(t)dt) = \mathop{\mathcal{P}}_{x_0}^{x}(1+A\,a(t)dt) = \exp\left[A\int_{x_0}^{x} a(t)dt\right] . \tag{7}$$

We note the interesting fact that (5) and (7) are exactly the same formulae as those applying when our system $Y' = A \cdot Y$ is reduced to one differential equation with one unknown function!

Next we observe that if $B(x)$ is a continuous and non-negative matrix function which is absolutely exponentiable in (x_0, x) and which for all values in the interval $x_0 \leq t \leq x$ satisfies

$$|A(t)| \leq B(t) \qquad \text{for all } t \text{ in } x_0 \leq t \leq x , \tag{8}$$

then it follows directly from the definition of the product-integral that both the product-integrals of A and of B exist and that

$$\left|\mathop{\mathcal{P}}_{x_0}^{x}(1+A(t)dt)\right| = \lim_{m\to\infty}\left|\prod_{i=0}^{m-1}(1+A(x_i)\Delta_i)\right| \leq$$

$$\lim_{m\to\infty}\prod_{i=0}^{m-1}(1+B(x_i)\Delta_i) = \mathop{\mathcal{P}}_{x_0}^{x}(1+B(t)dt) . \tag{9}$$

In particular it follows from (7) and (9) that

$$\left|\mathop{\mathcal{P}}_{x_0}^{x}(1+A(t)dt)\right| \leq \exp[K(x-x_0)] \tag{10}$$

in which

$$K = \max_{x_0 \leq t < x} |A(t)| \, . \tag{11}$$

We note, that we have already previously deduced (10), namely in (2.4.4).

Finally, we shall discuss two special cases in which the evaluation of the product-integral is still simpler than in the case (5).

Let the operator matrix $A(x)$ be a *column half matrix* (cf. § 7.3; type III § 7.6) which means that all the elements *below* the main diagonal vanish, i. e.

$$(n|A(x)|n') = 0 \quad \text{for all } n > n' \, . \tag{12}$$

It then follows from (2.4.6) and corollary 1 § 7.3 that the same is the case for $F_\nu(x, x_0)$, i. e.

$$(n|F_\nu(x, x_0)|n') = 0 \qquad \begin{array}{l} \text{for all } n > n' \\ \text{,, \ ,, } \nu = 0, 1, 2, \dots \\ \text{,, \ ,, } x \geq x_0 \, . \end{array} \tag{13}$$

From (2.4.1) and theorem 1 § 7.3 it finally follows that the product-integral exists and has also the same property

$$(n|F(x, x_0)|n') = \left(n \Big| \prod_{x_0}^{x} (1 + A(t)dt) \Big| n' \right) = 0 \qquad \begin{array}{l} \text{for all } n > n' \\ \text{,, \ ,, } x \geq x_0 \, . \end{array} \tag{14}$$

But from the fundamental equation $Y' = A \cdot Y$ it is then seen that each $(n|F(x, x_0)|n')$ will *for each fixed* n' depend only on the *higher* F-functions, i. e. $n > n'$, and due to (12) and (14) the fundamental equation, consequently, reads

$$\frac{\partial}{\partial x}(n'|F(x, x_0)|n') - (n'|A(x)|n')(n'|F(x, x_0)|n') = 0$$

$$\text{for all } n' = 0, 1, 2, \dots \tag{15}$$

$$\frac{\partial}{\partial x}(n|F(x, x_0)|n') - (n|A(x)|n)(n|F(x, x_0)|n') = \sum_{n''=n+1}^{n'} (n|A(x)|n'')(n''|F(x, x_0)|n')$$

$$\begin{array}{l} \text{for all } n' = 1, 2, 3, \dots \\ \text{,, \ ,, } n < n' \, . \end{array}$$

From the general theory of linear differential equations of the first order this system is known to have one and only one solution, which, due to the initial condition $F(x_0, x_0) = 1$, is given by

$$(n'|F(x, x_0)|n') = \exp\left[\int_{x_0}^{x} (n'|A(t)|n') \, dt \right]$$

$$\begin{array}{l} \text{for all } n' = 0, 1, 2, \dots \\ \text{,, \ ,, } x \geq x_0 \end{array}$$

$$(n|\boldsymbol{F}(x, x_0)|n') = \exp\left[\int_{x_0}^{x}(n|\boldsymbol{A}(t)|n)dt\right]\cdot \tag{16}$$

$$\left(\int_{x_0}^{x}\exp\left[-\int_{x_0}^{t}(n|\boldsymbol{A}(\tau)|n)d\tau\right]\sum_{n''=n+1}^{n'}(n|\boldsymbol{A}(t)|n'')(n''|\boldsymbol{F}(t, x_0)|n')dt\right)$$

for all $n' = 1, 2, 3, \ldots$
,, ,, $n < n'$
,, ,, $x \geq x_0$.

This system can be solved successively.

The second simple case we shall discuss is that in which the operator matrix $\boldsymbol{A}(x)$ is a *row half matrix* (cf. § 7.3; type III § 7.6) which means that all the elements *above* the main diagonal vanish, i. e.

$$(n|\boldsymbol{A}(x)|n') = 0 \qquad \text{for all } n < n' . \tag{17}$$

Then it follows from (2.4.6) and corollary 1 § 7.3 that we have also

$$(n|\boldsymbol{F}_\nu(x, x_0)|n') = 0 \qquad \text{for all } n < n' \tag{18}$$

,, ,, $\nu = 0, 1, 2, \ldots$
,, ,, $x \geq x_0$.

From (2.4.1) and theorem 1 § 7.3 it finally follows that the product-integral exists and has also the same property

$$(n|\boldsymbol{F}(x, x_0)|n') = \left(n|\prod_{x_0}^{x}(\boldsymbol{1} + \boldsymbol{A}(t)dt)|n'\right) = 0 \quad \text{for all } n < n' \tag{19}$$

,, ,, $x \geq x_0$.

This latter fact can, however, now be seen directly, in contrast to the previous case. From the fundamental equation $\boldsymbol{Y}' = \boldsymbol{A}\cdot\boldsymbol{Y}$ it is, namely, seen that (17) implies that each $(n|\boldsymbol{F}(x, x_0)|n')$ will *for each fixed* n' depend only on the *lower* $\boldsymbol{F}$-functions, i. e. $n < n'$, and the fundamental equation in analogy with (15), consequently, reads

$$\frac{\partial}{\partial x}(0|\boldsymbol{F}(x, x_0)|n') - (0|\boldsymbol{A}(x)|0)(0|\boldsymbol{F}(x, x_0)|n') = 0$$

for all $n' = 0, 1, 2, \ldots$

$$\frac{\partial}{\partial x}(n|\boldsymbol{F}(x, x_0)|n') - (n|\boldsymbol{A}(x)|n)(n|\boldsymbol{F}(x, x_0)|n') = \sum_{n''=0}^{n-1}(n|\boldsymbol{A}(x)|n'')(n''|\boldsymbol{F}(x, x_0)|n') \tag{20}$$

for all $n' = 0, 1, 2, \ldots$
,, ,, $n = 1, 2, 3, \ldots$.

This system has one and only one solution which, due to the initial condition $\boldsymbol{F}(x_0, x_0) = \boldsymbol{1}$, is given by the expressions, analogous to (16),

$$(n|\boldsymbol{F}(x, x_0)|n') = 0 \qquad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \ ,, \ n < n' \quad ,, \ ,, \ x \geq x_0$$

$$(n'|\boldsymbol{F}(x, x_0)|n') = \exp\left[\int_{x_0}^{x} (n'|\boldsymbol{A}(t)|n')dt\right] \qquad ,, \ ,, \ n' = 0, 1, 2, \ldots \quad ,, \ ,, \ x \geq x_0$$

$$(n|\boldsymbol{F}(x, x_0)|n') = \exp\left[\int_{x_0}^{x} (n|\boldsymbol{A}(t)|n)dt\right] \cdot \tag{21}$$

$$\left(\int_{x_0}^{x} \exp\left[-\int_{x_0}^{t} (n|\boldsymbol{A}(\tau)|n)d\tau\right] \sum_{n''=n'}^{n-1} (n|\boldsymbol{A}(t)|n'')(n''|\boldsymbol{F}(t, x_0)|n')dt\right)$$

$$\text{for all } n' = 0, 1, 2, \ldots \quad ,, \ ,, \ n > n' \quad ,, \ ,, \ x \geq x_0 .$$

This system can again be solved successively. We see that (21) agrees with (19). We, furthermore, note that when condition (17) is satisfied, the proof of the theorem of uniqueness is, thus, elementary.

§ 2.7. In the preceding paragraphs, §§ 2.2–2.6, we have shown that our fundamental equation (2.1.1a), $\boldsymbol{P}' = \boldsymbol{A} \cdot \boldsymbol{P}$, which governs our stochastic process, has one and only one solution in the interval $s \leq \tau \leq t$. Furthermore, we have seen that this solution is also the unique solution of the adjointed equation (2.1.2a) and satisfies the fundamental conditions (2.1.Ia)–(2.1.IIIa). The interval (s, t) is, however, here limited by the condition that the operator matrix $\boldsymbol{A} = (\boldsymbol{\Pi} - 1) \cdot \boldsymbol{p}$ shall be absolutely exponentiable in (s, t). We shall now show that this restriction is, as a rule, irrelevant and that the solution given by (2.5.7), i. e. the product-integral

$$\boldsymbol{P}(s, t) = \mathcal{P}_{s}^{t}(1 + \boldsymbol{A}(\tau)d\tau) = \lim_{'m' \to \infty} \prod_{i=0}^{m-1} (1 + \boldsymbol{A}(\tau_i)1_i) , \tag{1}$$

exists and is the solution for all values of s and t for which $\boldsymbol{A}$ is defined. Before we can do this we have, however, to investigate the conditions (2.1.IVa) and (2.1.Va), i. e. to show that (1) satisfies the two remaining fundamental conditions

$$(n|\boldsymbol{P}(t, s)|n') = P(n, t; n', s) \geq 0 \qquad \text{for all } n, n' = 0, 1, 2, \ldots \quad ,, \ ,, \ t \geq s \tag{2}$$

and

$$(\sum \cdot \boldsymbol{P}(t, s)|n') = \sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') = 1 \qquad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \ ,, \ t \geq s . \tag{3}$$

We note that the matrix appearing in the definition of the product-integral

$$\boldsymbol{T}(t, \Delta t) = \mathbf{1} + \boldsymbol{A}(t)\Delta t \tag{4}$$

has an 'anschaulich' interpretation as its (n, n') element, due to (2.1.12), simply denotes the asymptotic expression for the transition probability *from* the state n' *to* the state n, i. e. we have asymptotically

$$(n|\boldsymbol{P}(t+\Delta t, s)|n') = \sum_{n''=0}^{\infty}(n|\boldsymbol{T}(t, \Delta t)|n'')(n''|\boldsymbol{P}(t, s)|n') \,. \tag{5}$$

We shall, therefore, speak of $\boldsymbol{T}$ as the *transition probability operator.*

Now it is directly seen from (1) that (2) holds good. In the product on the right hand side of (1) we have, namely, from (2.1.12) that all the non-diagonal elements of each factor are non-negative and that the diagonal elements are of the form

$$(n|\mathbf{1}+\boldsymbol{A}(\tau_i)\Delta_i|n) = 1 - p(n, \tau_i)\Delta_i \qquad \text{for all } n = 0, 1, 2, \ldots \quad ,, \quad ,, \quad s \leqq \tau_i \leqq t \tag{6}$$

which expression becomes positive from a certain stage. The limit is, consequently, non-negative, q.e.d. We observe that (2) can, furthermore, be sharpened. In fact solving (2.1.1) with respect to $(n|\boldsymbol{P}|n')$ and using (2.1.5) and (2.1.1), we obtain

$$(n|\boldsymbol{P}(t, s)|n') = \exp\left[-\int_s^t p(n, \tau)d\tau\right] \cdot \tag{7}$$

$$\left(\int_s^t \exp\left[\int_s^\tau p(n, \tau')d\tau'\right]\sum_{n''=0}^{\infty}(n|\boldsymbol{\Pi}(\tau)\cdot\boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')d\tau + \delta_{n,n'}\right).$$

Due to (2.1.3), (2.1.4) and (2) all the factors on the right hand side of (7) are non-negative and if the integral is zero for some value t_0 of t, the integrand must, being a continuous function, be *identically* zero. (2) can thus be sharpened to

$$(n|\boldsymbol{P}(t, s)|n') \begin{cases} \text{either} = 0 & \text{for all } t \text{ in } s \leq t \leq t_0 \\ \quad\ \ \text{or} > 0 & \ ,, \quad ,, \quad t > s \end{cases} \tag{8}$$
$$\text{for all } n, n' = 0, 1, 2, \ldots \,.$$

The other condition (3) can, however, not be true for an *arbitrary* stochastic process. This fact is, besides, physically obvious as it is a priori possible that our stochastic variable can increase so strongly with the time or make so large jumps upwards that it can reach the value 'infinity' with a positive probability for a *finite* value of t.

As we must require the total 'probability mass' to be constant, equal to one, all what we can hope to prove generally, valid for an arbitrary stochastic process, is not the relation (3) but only the relation

$$\left(\sum \cdot \boldsymbol{P}(t, s)|n'\right) \leq 1 \qquad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geq s\,. \tag{9}$$

The missing amount we can here interpret as the probability of the value 'infinity', i. e.

$$P(\infty, t; n', s) = 1 - \left(\sum \cdot \boldsymbol{P}(t, s)|n'\right) \geq 0\,. \tag{10}$$

(9) is now easily proved from the fundamental equation (2.1.Ia). Integrating this equation from s to t, using the fact that each of the series defining $\boldsymbol{A} \cdot \boldsymbol{P}$ is *uniformly* convergent and summing over all n from 0 to N, keeping n' fixed, we obtain, due to (2.1.Ia),

$$\sum_{n=0}^{\infty}{}' \left(n|\boldsymbol{P}(t, s)|n'\right) = \lim_{N \to \infty} \sum_{n=0}^{N}{}' \left(n|\boldsymbol{P}(t, s)|n'\right) =$$

$$\Delta(N \geq n') + \lim_{N \to \infty} \sum_{n=0}^{N}{}' \sum_{n''=0}^{\infty}{}' \int_s^t \left(n|\boldsymbol{A}(\tau)|n''\right)\left(n''|\boldsymbol{P}(\tau, s)|n'\right) d\tau =$$

$$\Delta(N \geq n') + \lim_{N \to \infty} \sum_{n''=0}^{\infty}{}' \int_s^t \left(\sum_{n=0}^{N}{}' \left(n|\boldsymbol{A}(\tau)|n''\right)\right)\left(n''|\boldsymbol{P}(\tau, s)|n'\right) d\tau \tag{11}$$

in which the Δ-symbol is given in (1.2.10). Firstly, the inversion of the order of summation is here legitimate because we have simply a *finite* sum of convergent series. Secondly, the left hand side of (11) is, due to (8), monotonously increasing with N and the right hand side, therefore, either tends to a finite limit or to ∞. Now for each fixed value of n'' we have, due to (2.1.11), (2.1.6) and (8),

$$\int_s^t \left(\sum_{n=0}^{N}{}' \left(n|(\boldsymbol{\Pi}(\tau) - \boldsymbol{1}) \cdot \boldsymbol{p}(\tau)|n''\right)\right)\left(n''|\boldsymbol{P}(\tau, s)|n'\right) d\tau \leq 0 \quad \text{for all } N \geq n''\,. \tag{12}$$

Consequently, the right hand side of (11) is convergent and

$$\lim_{N \to \infty} \sum_{n''=0}^{\infty}{}' \int_s^t \left(\sum_{n=0}^{N}{}' \left(n|\boldsymbol{A}(\tau)|n''\right)\right)\left(n''|\boldsymbol{P}(\tau, s)|n'\right) d\tau \leq 0\,. \tag{13}$$

(13) inserted into (11) proves (9).

We are now able to show that (1) exists and is the solution in the whole region of definition of $\boldsymbol{A}$. Either we have that $\boldsymbol{A}$ is abso-

lutely exponentiable in any arbitrary interval, in which case there is nothing to show, or a 'radius of exponentiability' exists, i. e. a radius of convergence of the series (2.1.16). Let t_1 be a point within this radius. We can then start from this point and repeat the procedure in § 2.4, calculating a fundamental solution $\boldsymbol{P}(t, t_1)$ in a new interval $t_1 \leqq t \leqq t_2$. From (8), (9) and § 2.4 it is then obvious that the product

$$\boldsymbol{P}(t, s) = \boldsymbol{P}(t, t_1) \cdot \boldsymbol{P}(t_1, s) \tag{14}$$

exists and is a fundamental solution in the *whole* interval $s \leqq t \leqq t_2$. This procedure, which we shall call **exponentiable continuation** because it corresponds to the well-known procedure of *analytic* continuation, we may now repeat ad infinitum, obtaining a series of continuation points $t_0 = s < t_1 < t_2 < \ldots$ and we thus see that (1) represents in fact the solution in the *whole* region of definition of the operator matrix $\boldsymbol{A}(t)$ even if this matrix is only exponentiable in a *limited* interval.

Example 1.

It is, of course, possible that the exponentiable continuation stops *within* the region of definition of $\boldsymbol{A}(t)$. In such case our theory would turn out to be too narrow. We shall, however, not meet any such cases in our applications of the theory as our operator matrices, being defined in the whole region $s \leqq t < \infty$, will either be exponentiable in each interval (e. g. matrices of type I–III § 7.6) or the exponentiable continuation may be carried through to arbitrary high values (e. g. matrices of type IV § 7.6). As this latter case will be very important for our applications (cf. chap. 4) we shall consider it in further detail here. As shown in theorem 1 § 7.4 (or theorem 4 § 7.5 for m-dimensional matrices) matrices of type IV will be exponentiable in each interval limited by the condition (7.4.3) (or (7.5.38)). We shall later discuss stochastic processes with $f(t) = 2(\lambda + \gamma t)$ in which λ and γ are arbitrary non-negative constants (cf. [1]) p. 111, [1]) p. 117 and (7.1.8)). Thus one sequence of exponentiable continuation points will be given by the equation

$$t_n - t_{n-1} = \theta \frac{1}{2l(\lambda + \gamma t_n)} \qquad \begin{aligned} &0 < \theta < 1 \\ &t_0 = s = 0 \\ &n = 1, 2, 3, \ldots \end{aligned} \tag{15}$$

in which θ is an arbitrary number in the interval mentioned. Solving (15) we find

$$t_n = \frac{1}{2}\left(t_{n-1} - \frac{\lambda}{\gamma}\right) + \frac{1}{2}\left(\left(t_{n-1} + \frac{\lambda}{\gamma}\right)^2 + \frac{2\theta}{l\gamma}\right)^{\frac{1}{2}} =$$

$$t_{n-1} + \frac{1}{2}\left(t_{n-1} + \frac{\lambda}{\gamma}\right)\left(\left(1 + \frac{2\theta}{l\gamma\left(t_{n-1} + \frac{\lambda}{\gamma}\right)^2}\right)^{\frac{1}{2}} - 1\right) > t_{n-1} \,. \tag{16}$$

As a control we have from (16) that

$$t_n \underset{\gamma \to 0}{\to} t_{n-1} + \frac{\theta}{2l\lambda} \tag{17}$$

in agreement with (15). It is now easily seen that the monotonously increasing sequence (16) is unbounded also for $\gamma > 0$ and, consequently, any arbitrary point (> 0) can be reached by an exponentiable continuation. If, namely, the sequence were bounded, then it would have an upper limit, T, i. e. both $t_{n-1} \underset{n \to \infty}{\to} T$ and $t_n \underset{n \to \infty}{\to} T$. From (16) we would, however, obtain

$$t_n \underset{n \to \infty}{\to} T + \frac{1}{2}\left(T + \frac{\lambda}{\gamma}\right)\left(\left(1 + \frac{2\theta}{l\gamma\left(T + \frac{\lambda}{\gamma}\right)^2}\right)^{\frac{1}{2}} - 1\right) > T \tag{18}$$

which contradicts our assumption. Consequently, we have $t_n \to \infty$, q. e. d.

§ 2.8. Before in the next paragraph we discuss the conditions for the sign of equality to hold in (2.7.9) we shall discuss some simple types of stochastic processes.

The first case is that of a **finite stochastic process,** which is a process in which the operator matrix $\boldsymbol{A} = (\boldsymbol{\Pi} - \mathbf{1}) \cdot \boldsymbol{p}$ is finite of some order $N = 1, 2, 3, \ldots$ (cf. § 7.2; type II § 7.6), i. e.

$$(n|\boldsymbol{A}(t)|n') = 0 \quad \text{for all } n > N$$
$$\quad\quad\quad\quad\quad\quad\quad \text{,, \quad ,, } n' > N\,. \tag{1}$$

Such finite stochastic processes have first been discussed systematically by Kolmogoroff[1]) who proved that when $\boldsymbol{A}(t)$ is a given continuous finite probability operator, then one and only one solution of (2.1.1a) exists which, being also the solution of (2.1.2a), satisfies the fundamental conditions (2.1.Ia)-(2.1.Va). This case is now a special case of our general theory for, as shown in theorem 1 § 7.2, $\boldsymbol{A}$ is absolutely exponentiable in any interval. Furthermore, the solution is, of course, again finite of order N. We thus only have to prove that also (2.1.Va), or (2.7.3), is satisfied. This, however, is easily done, all summations being now finite and we thus have from (2.1.1a) and (2.1.13)

$$\sum_{n=0}^{N} \frac{\partial}{\partial t}(n|\boldsymbol{P}(t,s)|n') = \frac{\partial}{\partial t}\sum_{n=0}^{N}(n|\boldsymbol{P}(t,s)|n') =$$
$$\sum_{n=0}^{N}\sum_{n''=0}^{N}(n|\boldsymbol{A}|n'')(n''|\boldsymbol{P}|n') = \sum_{n''=0}^{N}\left(\sum_{n=0}^{N}(n|\boldsymbol{A}|n'')\right)(n''|\boldsymbol{P}|n') = 0\,. \tag{2}$$

[1]) Kolmogoroff (1931). Cf. p. 127.

(2) implies, because of the initial condition $\boldsymbol{P}(s,s) = \mathbf{1}$,

$$\sum_{n=0}^{N} (n|\boldsymbol{P}(t, s)|n') = \text{constant} = \sum_{n=0}^{N} (n|\boldsymbol{P}(s, s)|n') = 1 \qquad (3)$$
$$\text{for all } n' = 0, 1, 2, \ldots, N$$
$$\text{,, ,, } t \geqq s ,$$

q. e. d.

Another simple case in which (2.1.Va) can also be shown directly is that of a **pure absorption** or **'death' process**[1]). By this term we shall understand a stochastic process in which the stochastic variable can only *decrease* whenever it changes. This means that in the *relative transition probability matrix* $\boldsymbol{\Pi}(t)$ and, consequently, also in $\boldsymbol{A}(t)$, all elements *below* the main diagonal vanish, i. e.

$$(n|\boldsymbol{\Pi}(t)|n') = (n|\boldsymbol{A}(t)|n') = 0 \quad \text{for all } n > n' \qquad (4)$$
$$\text{,, ,, } t \geq s .$$

The operator matrix $\boldsymbol{A}$ is thus a **column half matrix** (cf. § 7.3; type III § 7.6) and this is just a special case of our general theory which we have discussed separately in § 2.6. As shown there the distribution matrix $\boldsymbol{P}(t, s)$ will then, of course, also be a column half matrix (cf. (2.6.14)), i. e.

$$(n|\boldsymbol{P}(t, s)|n') = 0 \quad \text{for all } n > n' \qquad (5)$$
$$\text{,, ,, } t \geq s .$$

Furthermore, it may be evaluated by successive quadratures (cf. (2.6.16)). We note that for each fixed value of $n' = 0, 1, 2, \ldots$ we may, without altering the stochastic process considered, assume $p(n, t) = 0$ for all $n > n'$. A *pure absorption* process can thus in a certain way be regarded as a special case of a *finite* process. The calculation (2)–(3), consequently, also holds good in this case if only we substitute n' for N. Thus (2.1.Va) is true for an arbitrary pure absorption process.

Another interesting case, which is *formally* analogous to the case of a pure *absorption* process, is that of a **pure propagation** or **'birth' process**[2]). By this term we shall understand a stochastic

[1]) Such processes may be met with e. g. in biology when one considers a family of individuals which can *die* but not *propagate*. Or in physics when one considers a number of radioactive atoms which in time can disintegrate, i. e. 'die'.

[2]) Such processes may be met with e. g. in biology when one considers a family of individuals which can *propagate* but not *die*. We shall later also meet them in physical problems, cf. chap 4. One also meets them in many other problems, e. g. in telephone

process in which the stochastic variable can only *increase* whenever it changes. This means that in the *relative transition probability matrix* $\boldsymbol{\Pi}(t)$ and, consequently, also in $\boldsymbol{A}(t)$, all elements *above* the main diagonal vanish, i. e.

$$\left(n|\boldsymbol{\Pi}(t)|n'\right) = \left(n|\boldsymbol{A}(t)|n'\right) = 0 \quad \text{for all } n < n' \quad ,, \quad ,, \quad t \geqq s \ . \tag{6}$$

The operator matrix is thus a **row half matrix** (cf. § 7.3; type III § 7.6) and this is just the second special case of our general theory which we have discussed separately in § 2.6. As shown there the distribution matrix $\boldsymbol{P}(t, s)$ will then, of course, also be a row half matrix (cf. (2.6.19)), i. e.

$$\left(n|\boldsymbol{P}(t, s)|n'\right) = 0 \quad \text{for all } n < n' \quad ,, \quad ,, \quad t \geqq s \ . \tag{7}$$

Furthermore, it may, just as for pure *absorption* processes, be evaluated by successive quadratures (cf. (2.6.21)). We note, however, that although the pure *propagation* processes are formally analogous to the pure *absorption* processes they are, nevertheless, essentially different as the first processes behave as *infinite* processes whereas the latter behave as *finite* processes. As we shall see in the next paragraph the calculation in (2) will no longer hold generally true for the pure propagation processes.

We shall, however, later on see that also for stochastic processes with operator matrices of type I, *type* II *and type* IV (*cf.* § 7.6) *the sign of equality holds in* (2.7.9), *i. e. the condition* (2.1.Va) *is fulfilled.*

Finally we shall consider the problem of the **absolute probability distribution.** If $\boldsymbol{P}(s)$ is an arbitrary matrix function consisting of only one column which satisfies

$$\left(n|\boldsymbol{P}(s)\right) = P(n, s) \geqq 0 \qquad \text{for all } n = 0, 1, 2, \ldots \tag{8}$$

and

$$\Sigma \cdot \boldsymbol{P}(s) = \sum_{n=0}^{\infty} P(n, s) = 1 \ , \tag{9}$$

then we can interpret $P(n, s)$ as the absolute probability of the stochastic variable assuming the value n at the time s. We now have

Theorem 1.

If $\mathbf{P}(s)$ *is an arbitrary matrix consisting of only one column representing the absolute probability distribution at the time* s, *i. e. satisfying* (8) *and* (9), *then*

theory when the stochastic variable denotes the number of calls within a certain period. Or in actuarial problems, cf. Lundberg (1940) and so on.

$$\boldsymbol{P}(t) = \boldsymbol{P}(t, s) \cdot \boldsymbol{P}(s) \,, \tag{10}$$

in which $\boldsymbol{P}(t, s)$ *is the relative distribution given by the product-integral* (2.7.1), *exists, is a solution of* (2.1.Ia) *and represents the absolute probability distribution at the time* t, *i. e. satisfies*

$$\lim_{t \to s} \boldsymbol{P}(t) = \boldsymbol{P}(s) \,, \tag{I}$$

$$\boldsymbol{P}(t) \geqq 0 \qquad \text{for all } t \geqq s \tag{II}$$

and

$$\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t) \leqq 1 \qquad \text{,, \quad ,, } t \geqq s \,. \tag{III}$$

From (8), (9), (2.7.2) and (2.7.9) it follows at once that $\boldsymbol{P}(t)$ given by (10) exists. From § 2.4 it follows that (10) is a solution of (2.1.Ia) and satisfies (I), all the sums in (10) being uniformly convergent in $s \leqq \tau \leqq t$. (II) follows directly from (2.7.2) and (8), whereas (III) follows from (2.7.9) and (9) because

$$\sum_{n=0}^{\infty} (n|\boldsymbol{P}(t)) = \sum_{n=0}^{\infty} \sum_{n'=0}^{\infty} (n|\boldsymbol{P}(t, s)|n')(n'|\boldsymbol{P}(s)) =$$

$$\sum_{n'=0}^{\infty} \left(\sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') \right) (n'|\boldsymbol{P}(s)) \leqq \sum_{n'=0}^{\infty} (n'|\boldsymbol{P}(s)) = 1 \,. \tag{11}$$

Here the inversion of the order of summation is legitimate because all the terms are non-negative. We note that if the sign of equality holds in (2.7.9), i. e. (2.1.Va) is valid, then it also holds in (11) and if it does not hold for only a single value of n', then it does not hold in (11) either.

§ 2.9. We shall now discuss the main problem of the theory, namely the conditions for the sign of equality to hold in (2.7.9), i. e. for (2.1.Va) to be satisfied:

$$(\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s)|n') = 1 \qquad \text{for all } n' = 0, 1, 2, \ldots \quad \text{,, \quad ,, } t \geqq s \,. \tag{1}$$

Feller[1]) was the first to suggest such conditions and Lundberg[2]) has proved them in the case of **elementary stochastic processes,** i. e. pure propagation processes in which the stochastic variable can change by only $+1$ whenever it changes:

$$(n|\boldsymbol{\Pi}(t)|n') = f(n', t)\delta_{n, n'+1} \qquad \text{for all } n' = 0, 1, 2, \ldots \,. \tag{2}$$

From (2) it follows that $\boldsymbol{\Pi}(t)$ and, consequently, also the operator matrix $\boldsymbol{A}(t)$, of an elementary stochastic process is both a *row half matrix* and a *column semi-diagonal matrix with* $l = 1$ (cf. § 7.3) and

[1]) Feller (1939) p. 21.

[2]) Lundberg (1940) p. 41 ff.

this process is thus a special case within our general theory.

We shall now investigate the conditions in question from the point of view of the general stochastic processes which we are considering. We start by quoting the

Feller-Lundberg theorem.

(I) *For elementary stochastic processes, characterized by* (2), *a* **sufficient** *condition for the validity of* (1) *in the whole interval* $s \leq \tau \leq t$ *consists in the divergence of the series*

$$\sum_{n=0}^{\infty} \frac{1}{\bar{p}(n)} = \infty \tag{3}$$

in which

$$\bar{p}(n) = \max_{s \leq \tau \leq t} p(n, \tau) . \tag{4}$$

(II) *For elementary stochastic processes, characterized by* (2), *a* **necessary** *condition for the validity of* (1) *in the whole interval* $s \leq \tau \leq t$ *consists in the divergence of the series*

$$\sum_{n=0}^{\infty} \frac{1}{\underline{p}(n)} = \infty \tag{5}$$

in which

$$\underline{p}(n) = \min_{s \leq \tau \leq t} p(n, \tau) . \tag{6}$$

We first note that the contents of the two conditions (I) and (II) consist in keeping the speed of propagation, the 'birth rate', within reasonable limits, i. e. assuring that the intensity function $p(n, t)$ does not increase too strongly with n. It is quite plausible that this must be a condition for the validity of (1) since, as already mentioned (cf. § 2.7), if the intensity function increased too strongly, it would cause a certain 'probability mass' to 'flow out' into the point 'infinity' in the course of a *finite* time, i. e. our stochastic variable would, for a finite value of t, become a *positive* probability for assuming the value 'infinity' (cf. (2.7.10)). As no 'probability mass' can be absorbed or created the *total* 'probability mass' must be constantly equal to one. This fact would make the sum in (1), representing the probability for all *finite* values, *less* than one.

These 'anschaulich' considerations can also be expressed purely analytically. Integrating our fundamental equation (2.1.1a) from s to t gives, using (2.1.Ia),

$$1 - \boldsymbol{P}(t, s) = \int_s^t \boldsymbol{p}(\tau) \cdot \boldsymbol{P}(\tau, s) d\tau - \int_s^t \Pi(\tau) \cdot \dot{\boldsymbol{p}}(\tau) \cdot \boldsymbol{P}(\tau, s) d\tau \tag{7}$$

and next we obtain by summation, using (2.1.6) and (2.7.9),

$$0 \leq 1 - \sum_{n=0}^{N} (n|\boldsymbol{P}(t, s)|n') = \sum_{n''=0}^{N} \int_s^t (n''|\boldsymbol{p}(\tau) \cdot \boldsymbol{P}(\tau, s)|n')d\tau -$$

$$\sum_{n=0}^{N} \sum_{n''=0}^{\infty} \int_s^t (n|\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')d\tau =$$

$$\left(\sum_{n=0}^{\infty} \sum_{n''=0}^{N} - \sum_{n=0}^{N} \sum_{n''=0}^{\infty}\right) \int_s^t (n|\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')d\tau =$$

$$\left(\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} - \sum_{n=0}^{N} \sum_{n''=N+1}^{\infty}\right) \int_s^t (n|\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')d\tau =$$

$$\text{'up'—'down'} < 1 \qquad \text{for all } n' = 0, 1, 2, \ldots;\quad N \geq n';\quad t \geq s. \tag{8}$$

We note that on the right hand side of (8) we have effectively *less* than one due to (2.7.8). In fig. 1 we have illustrated the operations
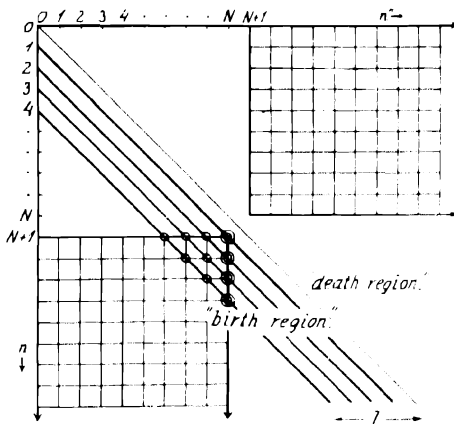
FIG. 1.

in (8). In this figure each point (n, n'') denotes the transition *from* the state n'' *to* the state n. If $n < n''$ the variable has decreased, i. e. a 'death' process, whereas $n > n''$ means that the variable has increased, i. e. a 'birth' process in our biological picture (cf. § 2.8). The upper double hatched region corresponds to $\sum_{n=0}^{N} \sum_{n''=N+1}^{\infty}$

whereas the lower one to $\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N}$. (For the rest of the figure, see later on.) The relation (8) could, however, have been written down at once as it has an immediate 'anschaulich' interpretation representing simply a *continuity equation* for the 'probability mass'. The left hand side gives, namely, the amount with which the total 'probability mass', corresponding to the states $n = 0, 1, 2, \ldots, N$, has decreased in the time $t-s$. And the right hand side gives the total 'probability current' which during the same time has passed the state $n = N$, the first term giving the total 'current' which has 'flowed upwards' *from* the states $n = 0, 1, 2, \ldots, N$ *to* the higher states while the second term gives the corresponding total downward 'current'.

We observe that this 'anschaulich' interpretation of the two terms in (8) justifies our previous remark that $1 - \sum_{n=0}^{\infty} (n|\boldsymbol{P}|n')$ may be interpreted as the probability of the stochastic variable reaching the value 'infinity' in a finite time. In fact the 'probability mass' which has 'flowed out' to the point 'infinity' must be defined as

$$P(\infty, t; n', s) = \lim_{N\to\infty} \left(\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} - \sum_{n=0}^{N} \sum_{n''=N+1}^{\infty} \right) \cdot$$

$$\int_s^t (n|\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')d\tau = 1 - \sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') \geq 0 \quad (9)$$

in which we have used (2.7.9). (9) is actually (2.7.10), q.e.d. We note, however, that although we have called the two terms on the right hand side of (8) 'probability masses' we cannot in general conclude that each of them will not exceed one, as it is possible that, so to say, the same 'mass' may pass the state $n = N$ several times in the time $t-s$.

Before proceeding we note that it is obvious that the FELLER-LUNDBERG theorem cannot be upheld generally. If e. g. $p(n, t) = 0$ for some value of n, n_0, (3) is obviously satisfied. If the process belongs to LUNDBERG's elementary processes, characterized by (2), this state $n = n_0$ cannot be surpassed because if the variable jumps into this 'hole' at some time it must remain there for all time. The process is, therefore, stopped altogether, i. e. it is a *finite* process and as proved in (2.8.3) the FELLER-LUNDBERG theorem is in fact true in this case. But in the general case, where the variable can jump by more than one step in a change, such a 'dead' state may simply be skipped, and it is quite possible that the variable may, in spite of the existence

of one or more such 'dead' states, reach the value '∞' within a finite time. (3) cannot, consequently, be a *sufficient* condition in the general case. Furthermore, we see that the necessary modifications must go in such a direction that if we have so many successive 'dead' states that they cannot be skipped in a single change, which fact would again stop the process, then the condition must be satisfied.

From (2.7.8) it now follows that the expression on the right hand side of (8) will, for fixed value of n', t and s, constitute the element of a *monotonous non-increasing positive sequence*. But such a sequence has a lower limit $\delta \geq 0$ for $N \to \infty$, i. e. we have from (8), (9) and the fact that both terms on the right of (8) are non-negative.

$$0 \leq \delta = 1 - \sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') = P(\infty, t; n', s) \leq$$
$$\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} \int_s^t (n|\boldsymbol{\Pi}(\tau)|n'')p(n'', \tau)(n''|\boldsymbol{P}(\tau, s)|n')d\tau \tag{10}$$

for all $n' = 0, 1, 2 \ldots$
" " $N \geq n'$
" " $t \geq s$.

In the case of the LUNDBERG elementary processes, characterized by (2), the second term in (8) vanishes and (10) thus reduces to, cf. fig. 1,

$$0 \leq \delta = 1 - \sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') = P(\infty, t; n', s) =$$
$$\int_s^t p(N, \tau)(N|\boldsymbol{P}(\tau, s)|n')d\tau \leq \overline{p}(N) \int_s^t (N|\boldsymbol{P}(\tau, s)|n')\, d\tau \tag{11}$$

in which $\overline{p}(N)$ is given in (4). From (11) LUNDBERG obtains, dividing on both sides with $\overline{p}(N)$ and summing over all $N \geq n'$, using (2.7.9),

$$\delta \sum_{N=n'}^{\infty} \frac{1}{\overline{p}(N)} \leq \int_s^t \sum_{N=n'}^{\infty} (N|\boldsymbol{P}(\tau, s)|n')\, d\tau \leq t - s < \infty \tag{12}$$

for all $n' = 0, 1, 2, \ldots$
" " $t \geq s$.

This relation can, however, only be true under the condition (3) if $\delta = 0$ and (3) is, consequently, *sufficient* to ensure the validity of (1) for *elementary processes*, q. e. d.

We observe that without altering LUNDBERG's proof we can generalize his notion of elementary processes so that it embraces also *absorption*, i. e. define a process as *elementary* if

$$(n|\boldsymbol{\Pi}(t)|n') \equiv 0 \quad \text{for all } n > n' + 1 \tag{13}$$

" " $n' = 0, 1, 2, \ldots$.

In other words (13) means that we give up the assumption that $\boldsymbol{\Pi}$, and thus also $\boldsymbol{A}$, is a *row half matrix* and only retain the assumption that they are *column semi-diagonal with $l = 1$* (cf. § 7.3).

We shall not here investigate the most general modification of the FELLER–LUNDBERG condition (3) applicable to the case of our general stochastic processes. We shall confine ourselves to the consideration of a case which will–presumably–cover all cases one can meet with in the practical applications of the theory. We shall, namely, assume in the following that *in each* **upward** *change our stochastic variable cannot change by more than a certain number of stages, say l, independent of the value of n' before the change*, i. e. the process allows at most of l-double 'births'

$$\left(n|\boldsymbol{\Pi}(t)|n'\right) = 0 \quad \text{for all } n > n'+l \qquad \text{,, ,, } n' = 0, 1, 2, \ldots \quad . \tag{14}$$

(14) means that we assume $\boldsymbol{\Pi}(t)$ and, consequently, also $\boldsymbol{A}(t)$, to be an arbitrary *column semi-diagonal* matrix, whereas the simple elementary processes given in (13) correspond to $l = 1$[1]. As (14) is a generalization of the elementary processes (2), or in the generalized sense (13), we shall call processes characterized by (14) **generalized elementary processes.** We see that (14) means that the relative transition matrix $\boldsymbol{\Pi}(t)$ and, consequently, also the operator matrix $\boldsymbol{A}(t)$, has in each column at most l non-vanishing elements *below* the main diagonal, cf. fig. 1 p. 59, but that we make no assumptions about the number of non-vanishing elements *above* the main diagonal, as these elements, which correspond to absorption or 'death' while the elements below correspond to propagation or 'birth', do not enter into (10).

For the generalized elementary processes we now have that (10), due to (2.1.6), reduces to the following expression, cf. fig. 1 p. 59 where the points denoted by single and double circles are the only ones entering into the sum on the right hand side of (10)[2],

[1]) We stress that the theory is, in our form, *not* applicable to an arbitrary operator matrix of this type as shown in example 1 § 7.3. Presumably, most processes met with in practice will, however, have operator matrices covered by the types I–IV (§ 7.6). We shall, consequently, in the following tacitly assume that $\boldsymbol{A}(t)$, besides being column semi-diagonal, is of such nature as to be covered by our theory.

[2]) We note that in $\sum_{n''=N-l+1}^{N}$ negative values of n'' may occur, but then we, of course, define $(n''|\boldsymbol{P}|n') = 0$ for $n'' < 0$.

$$0 \leq \delta = 1 - \sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s)|n') = P(\infty, t; n', s) \leq$$

$$\sum_{n''=N-l+1}^{N} \int_s^t p(n'', \tau)(n''|\boldsymbol{P}(\tau, s)|n')d\tau \leq \bar{p}(N) \sum_{n''=N-l+1}^{N} \int_s^t (n''|\boldsymbol{P}(\tau, s)|n')d\tau \tag{15}$$

for all $n' = 0, 1, 2, \ldots$
" " $N \geq n'$
" " $t \geq s$.

In (15) we have now defined, instead of (4),

$$\bar{p}(n) = \max_{s \leq \tau \leq t} \{p(n', \tau)\} \quad \text{for } n' = n-l+1, n-l+2, \ldots, n\,. \tag{16}$$

We see that with this new definition we can repeat LUNDBERG's argument (12). In fact we obtain from (15), using again (2.7.9),

$$\delta \sum_{N=n'}^{\infty} \frac{1}{\bar{p}(N)} \leq \sum_{N=n'}^{\infty} \sum_{n''=N-l+1}^{N} \int_s^t (n''|\boldsymbol{P}(\tau, s)|n')d\tau =$$

$$\int_s^t \sum_{n''=0}^{l-1} \sum_{N=n'}^{\infty} (N-n''|\boldsymbol{P}(\tau, s)|n')d\tau \leq l(t-s) < \infty \tag{17}$$

for all $n' = 0, 1, 2, \ldots$
" " $t \geq s$.

(17) shows that we have proved

Theorem 1.

For generalized elementary stochastic processes, characterized by (14) *i. e. having column semi-diagonal operator matrices, a* **sufficient** *condition for the validity of* (1) *in the whole interval $s \leq \tau \leq t$ consists in the divergence of the series*

$$\sum_{n=0}^{\infty} \frac{1}{\bar{p}(n)} = \infty \tag{18}$$

in which $\bar{p}(n)$ is given in (16).

As corollaries of theorem 1 we note the following ones, often met with in practice.

Corollary 1.

If the maximum values of $p(n, \tau)$, in the interval $s \leq \tau \leq t$, of a generalized elementary stochastic process, characterized by (14), *consti-*

tute a monotonous non-decreasing sequence, then the unmodified FELLER-LUNDBERG *theorem* (3) *is valid also for this process.*

This corollary follows at once from (16).

Corollary 2.

If for a generalized elementary stochastic process, characterized by (14), $\bar{p}(n)$ *given in* (16) *satisfies the condition*

$$\bar{p}(n) \leqq c \cdot n \quad \text{for all } n \geqq N \tag{19}$$

in which $c > 0$ *is an arbitrary constant and* $N > 0$ *is an arbitrary integer, then* (18) *is fulfilled and* (1) *consequently valid.*

In fact we have in this case

$$\sum_{n=N}^{\infty} \frac{1}{\bar{p}(n)} \geqq \frac{1}{c} \sum_{n=N}^{\infty} \frac{1}{n} = \infty \tag{20}$$

due to the divergence of the harmonious series. (19) shows, together with (7.1.8), that the operator matrix is of type IV (cf. § 7.6) and we thus see that (1) is valid in this case.

Corollary 3.

If for a generalized elementary stochastic process, characterized by (14), $\bar{p}(n)$ *given in* (16) *satisfies the condition*

$$\bar{p}(n) \leqq c \quad \text{for all } n = 0, 1, 2, \ldots \tag{21}$$

in which $c > 0$ *is an arbitrary constant, then* (18) *is fulfilled and* (1) *consequently valid.*

In fact we have in this case

$$\sum_{n=0}^{\infty} \frac{1}{\bar{p}(n)} \geqq \frac{1}{c} \sum_{n=0}^{\infty} \frac{1}{1} = \infty . \tag{22}$$

(21) shows, together with (7.1.8), that the operator matrix is *bounded*, i. e. of type I (cf. § 7.6), and we thus see that (1) is in fact valid for the stochastic processes considered by FELLER[1]), these processes thus being special cases of our general theory.

We note, however, that in FELLER's theory the same fact is shown to be true *without* the restrictive assumption of the processes being generalized elementary processes as given by (14). We could, by the way, also prove it

[1]) FELLER (1937).

generally here by utilizing the majorization given by (2.4.4) and (7.1.15) to show that the calculations in (2.8.2)–(2.8.3) are also valid for **arbitrary** *bounded* matrices.

Corollary 4.

If for a generalized elementary stochastic process, characterized by (14), *we have from a certain stage* n_0

$$p(n, t) \equiv 0 \quad \text{for all } n \geqq n_0\,, \tag{23}$$

then (18) *is satisfied and* (1) *consequently valid.*

This corollary follows at once from theorem 1. (23) shows, however, together with (14) that the operator matrix is in this case simply *finite* of an order at most equal to n_0+l and this corollary has thus already been proved direct in § 2.8. We note, however, that in corollary 4 the assumption of the process being a *generalized elementary* is unnecessary. In fact it follows direct from (10), inserting (23) and using (2.1.6), that $\delta = 0$. We thus see that (1) is valid for *arbitrary* stochastic processes with operator matrices of type II (cf. § 7.6).

Finally we observe that as regards stochastic processes with operator matrices of type III (cf. § 7.6) we have shown in § 2.8 that (1) is valid for *arbitrary* pure *absorption* processes. As shown in the next paragraph the same fact is, however, *not* generally true for pure *propagation* processes.

§ 2.10. We shall now shortly discuss the modifications of the second part of the FELLER-LUNDBERG theorem, (2.9.5)–(2.9.6), which are necessary when we generalize from the *elementary* processes (2.9.2) to our *generalized elementary* processes (2.9.14). For this purpose we consider again (2.9.8).

If we first apply it to the elementary processes, characterized by (2.9.2), it will be reduced to, cf. fig. 1 p. 59 and (2.9.11),

$$0 \leqq \int_s^t p(N, \tau)(N|\boldsymbol{P}(\tau, s)|n')\,d\tau < 1 \quad \begin{array}{l} \text{for all } n' = 0, 1, 2, \ldots \\ \;,, \quad ,, \quad N \geqq n' \\ \;,, \quad ,, \quad t \geqq s\,. \end{array} \tag{1}$$

Dividing by $p(N)$, given in (2.9.6), and summing over all $N \geqq n'$ LUNDBERG obtains

$$\int_s^t \sum_{N=n'}^{\infty} (N|\boldsymbol{P}(\tau, s)|n')\,d\tau < \sum_{N=n'}^{\infty} \frac{1}{p(N)}\,. \tag{2}$$

If now the series on the right hand side of (2) were convergent, then for each

fixed value of s and t we could find a value n_0' so that

$$\sum_{N=n'}^{\infty} \frac{1}{\underline{p}(N)} < t-s \quad \text{for all } n' \geqq n_0'(t, s) \qquad (3)$$
$$\text{,, each fixed } t > s\,.$$

Assuming (2.9.1) to be valid for all $n' = 0, 1, 2, \ldots$ we should thus, due to (2.8.7) which is applicable in this case as an elementary process is a special case of a pure propagation process, obtain the contradictory statement

$$t-s = \int_s^t \sum_{N=0}^{\infty} (N|\boldsymbol{P}(\tau, s)|n')\,d\tau = \int_s^t \sum_{N=n'}^{\infty} (N|\boldsymbol{P}(\tau, s)|n')\,d\tau < t-s \qquad (4)$$
$$\text{for all } n' \geqq n_0'(t, s)$$
$$\text{,, each fixed } t > s\,.$$

Our assumption (3) must, consequently, be wrong and the second part of the Feller-Lundberg theorem (2.9.5)-(2.9.6) is thus proved.

We now see that this argument is not so easily generalized as was the case with the *sufficient* condition (2.9.3). This fact is, however, not particularly deplorable as in the practical applications of the theory we are always far more interested in *sufficient* than in *necessary* conditions. We shall, therefore, not investigate the necessary modifications in the general case but content ourselves with proving that for an **arbitrary**, and not only *generalized elementary, pure propagation process*, characterized by (2.8.6) i.e. whose operator matrix is a *row half matrix*, the Feller-Lundberg condition (2.9.5)-(2.9.6) is still valid. In fact (2.9.8) will now, using (2.1.5), (2.1.6) and (2.8.6), cf. fig. 1 p. 59 in which the points denoted by double circles are the only points we keep in $\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N}$, be reduced to

$$0 \leqq \int_s^t p(N, \tau)(N|\boldsymbol{P}(\tau, s)|n')\,d\tau - \sum_{n=N+1}^{\infty} \int_s^t (n|\boldsymbol{\Pi}(\tau)|N)\,p(N, \tau)(N|\boldsymbol{P}(\tau, s)|n')\,d\tau \leqq$$
$$\sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} \int_s^t (n|\boldsymbol{\Pi}(\tau)\cdot \boldsymbol{p}(\tau)|n'')(n''|\boldsymbol{P}(\tau, s)|n')\,d\tau < 1 \qquad (5)$$
$$\text{for all } n' = 0, 1, 2, \ldots$$
$$\text{,, ,, } N \geqq n'$$
$$\text{,, ,, } t \geqq s\,.$$

(5) is, however, equivalent with (1) and the argument (2)-(4) can thus be repeated word for word and we have, consequently, proved

Theorem 1.

For an **arbitrary,** *and not necessarily generalized elementary, pure propagation process, characterized by* (2.8.6), *a* **necessary** *condition for the validity of* (2.9.1) *in the whole interval* $s \leq \tau \leq t$ *consists in the divergence of the series*

$$\sum_{n=0}^{\infty} \frac{1}{\underline{p}(n)} = \infty \tag{6}$$

in which

$$\underline{p}(n) = \min_{s \leq \tau \leq t} p(n, \tau) . \tag{7}$$

We finally note the following theorem, which is due to LUNDBERG, and which is 'anschaulich' obvious.

Theorem 2.

If for some value of n', *say* n'_0, *and of* (t, s), *say* (t_0, s_0), *we have for a pure propagation process*

$$\sum_{n=0}^{\infty} (n|\boldsymbol{P}(t_0, s_0)|n'_0) = \sum_{n=n'_0}^{\infty} (n|\boldsymbol{P}(t_0, s_0)|n'_0) < 1 , \tag{8}$$

then (8) *will hold good for all later times* $t \geq t_0$:

$$\sum_{n=0}^{\infty} (n|\boldsymbol{P}(t, s_0)|n'_0) < 1 \quad \text{for all } t \geq t_0 . \tag{9}$$

In fact, differentiating (2.9.8) with respect to t and using the fact that $\boldsymbol{\Pi}$, $\boldsymbol{p}$ and $\boldsymbol{P}$ are all non-negative, we have in this case

$$\frac{\partial}{\partial t} \sum_{n=0}^{N} (n|\boldsymbol{P}(t, s)|n') = - \sum_{n=N+1}^{\infty} \sum_{n''=0}^{N} (n|\boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t)|n'')(n''|\boldsymbol{P}(t, s)|n') \leq 0$$

$$\text{for all } n' = 0, 1, 2, \ldots \quad \text{,, ,, } N = 0, 1, 2, \ldots \quad \text{,, ,, } t \geq s . \tag{10}$$

(10) shows that for all $N = 0, 1, 2, \ldots$ the sum on the left hand side is a monotonous, non-increasing function of t which fact is, of course, also obvious from the 'anschaulich' interpretation in (2.9.8), the downward 'probability current' being now zero. The same fact must, therefore, be true for the convergent sum $\sum_{n=0}^{\infty} (n|\boldsymbol{P}|n')$ which fact together with (8) proves (9).

§ 2.11. A question which is often of interest in the practical applications of the theory is that of the existence and computation of the *moments*, both of the *relative* and of the *absolute* probability distributions. The **moment of order k** of the *relative* distribution $\boldsymbol{P}(t, s)$ is defined as the mean value of n^k:

$$\overline{n^k} = \overline{n^k}(t, s, n') = \sum_{n=0}^{\infty} n^k(n|\boldsymbol{P}(t, s)|n') \geqq 0 \tag{1}$$

$$\text{for all } n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s \quad ,, \quad ,, \quad k \geqq 0$$

in which k is an arbitrary non-negative constant (which is usually an integer). In the same way the moments of the *absolute* distribution $\boldsymbol{P}(t)$ are defined by

$$\overline{n^k} = \overline{n^k}(t) = \sum_{n=0}^{\infty} n^k(n|\boldsymbol{P}(t)) \geqq 0 \qquad \text{for all } t \geqq s \quad ,, \quad ,, \quad k \geqq 0. \tag{2}$$

Between (1) and (2) we have, of course, an intimate connection due to theorem 1 § 2.8. In fact, introducing (2.8.10) into (2) we have, as all summations are absolutely convergent because all terms entering are non-negative,

$$\overline{n^k}(t) = \sum_{n=0}^{\infty}\sum_{n'=0}^{\infty} n^k(n|\boldsymbol{P}(t, s)|n')(n'|\boldsymbol{P}(s)) =$$

$$\sum_{n'=0}^{\infty}\left(\sum_{n=0}^{\infty} n^k(n|\boldsymbol{P}(t, s)|n')\right)(n'|\boldsymbol{P}(s)) = \sum_{n'=0}^{\infty} \overline{n^k}(t, s, n')(n'|\boldsymbol{P}(s)). \tag{3}$$

Due to (3) it is, therefore, sufficient to investigate the moments of the *relative* distribution $\boldsymbol{P}(t, s)$.

Firstly, we note that for $k = 0$ the question is already settled as we have from (2.7.9)

$$\overline{n^0}(t, s, n') = (\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s)|n') \leqq 1 \quad \text{for all } n' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s \tag{4}$$

and from (2.8.III)

$$\overline{n^0}(t) = \boldsymbol{\Sigma} \cdot \boldsymbol{P}(t) \leqq 1 \qquad \text{for all } t \geqq s. \tag{5}$$

Secondly, we note that we have for $t = s$, due to $\boldsymbol{P}(s, s) = \boldsymbol{1}$,

$$\overline{n^k}(s, s, n') = n'^k \qquad \text{for all } n' = 0, 1, 2, \ldots. \tag{6}$$

As in § 2.9 we shall confine ourselves to the consideration of the

generalized elementary processes characterized by (2.9.14), i.e. whose operator matrices are *column semi-diagonal*[1]). We shall then prove the following theorem which is a natural extension of theorem 1 § 7.4.

Theorem 1.

For a generalized elementary stochastic process whose intensity function satisfies

$$p(n, t) \leqq f(t) \cdot n \quad \text{for all } n \geqq 1 \ , \tag{7}$$

i. e. whose operator matrix is of type IV (*cf.* § 7.6), *we have for all* $k \geqq 0$ *that the moments exist and are continuous in each interval* (s, t) *for which*

$$t - s < \frac{1}{lC} \tag{8}$$

with

$$C = 2 \max_{s \leqq \tau \leqq t} f(\tau) \ . \tag{9}$$

In fact we have, as a generalization of (7.4.5), using (2.4.4), (7.1.8), (7.3.8), (7), (9) and lemma 3 § 7.3,

$$\overline{n^k}(\tau, s, n') \leqq \sum_{n=0}^{\infty} n^k \big(n \,|\, \mathbf{exp}[\boldsymbol{K}(t-s)] \,|\, n'\big) = \sum_{\nu=0}^{\infty} \left(\sum_{n=0}^{n'+\nu l} n^k (n \,|\, \boldsymbol{K}^\nu \,|\, n') \right) \frac{(t-s)^\nu}{\nu !} \leqq$$

$$\sum_{\nu=0}^{\infty} (n' + \nu l)^k M_{n'} \left(\prod_{\beta=1}^{\nu-1} (n' + \beta l) \right) C^{\nu-1} \frac{(t-s)^\nu}{\nu !} = \sum_{\nu=0}^{\infty} v_\nu \tag{10}$$

for all $n' = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, τ in $s \leqq \tau \leqq t$
,, ,, $k \geqq 0$.

This series is, however, convergent for all values of $k \geqq 0$ and all intervals satisfying (8) because

$$\frac{v_{\nu+1}}{v_\nu} = \left(\frac{n' + (\nu+1) l}{n' + \nu l} \right)^k \frac{n' + \nu l}{\nu + 1} C(t-s) \underset{\nu \to \infty}{\rightarrow} lC(t-s) < 1 \quad \text{for all } k \geqq 0 \ . \tag{11}$$

Thus, all the moments exist, and as (10) shows that the convergence is uniform in $s \leqq \tau \leqq t$ they are, consequently, continuous, q. e. d.

(11) shows still more, namely that the *radius of convergence* of the series (10) is *independent of* k, being the same for all the moments.

---

[1]) We note that for *finite* and for *pure absorption* processes all the moments exist, of course, as all summations are finite.

We can thus use exactly the **same** *exponentiable continuation* as we used for $\boldsymbol{P}(t, s)$ itself (cf. § 2.7). We thus find that the restriction (8) is, in general, artificial, *the moments existing throughout the region of definition of the operator matrix.*

Next, we prove that also for $\boldsymbol{P}'$ all the moments exist. In fact we have by the same argument, using the fundamental equation (2.1.1a),

$$\left|\sum_{n=0}^{\infty}{}' n^k \left(n\left|\frac{\partial}{\partial\tau}\boldsymbol{P}(\tau, s)\right|n'\right)\right| = \left|\sum_{n=0}^{\infty}{}' n^k (n|\boldsymbol{A}(\tau)\cdot\boldsymbol{P}(\tau, s)|n')\right| \leq$$

$$\sum_{n=0}^{\infty}{}' n^k (n|\boldsymbol{K}\cdot\exp[\boldsymbol{K}(t-s)]|n') = \sum_{\nu=0}^{\infty}\left(\sum_{n=0}^{n'+(\nu+1)l}{}' n^k (n|\boldsymbol{K}^{\nu+1}|n')\right)\frac{(t-s)^\nu}{\nu!} \leq$$

$$\sum_{\nu=0}^{\infty}{}' (n'+(\nu+1)l)^k M_{n'}\left(\prod_{\beta=1}^{\nu}(n'+\beta l)\right) C^\nu \frac{(t-s)^\nu}{\nu!} \tag{12}$$

for all $n' = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, τ in $s \leq \tau \leq t$
,, ,, $k \geq 0$.

This series is, of course, again uniformly convergent in $s \leq \tau \leq t$ for all $k \geq 0$ with the same radius of convergence as the series in (10). This fact implies that the series on the left hand side of (12) is a continuous function. It may, furthermore, be integrated term by term with the result that, using (6),

$$\int_s^t \sum_{n=0}^{\infty}{}' n^k \left(n\left|\frac{\partial}{\partial\tau}\boldsymbol{P}(\tau, s)\right|n'\right)d\tau = \overline{n^k}(t, s, n') - n'^k . \tag{13}$$

As an integral of a continuous function is differentiable everywhere with a continuous differential coefficient we finally obtain that $\overline{n^k}$ is a differentiable function with respect to t with the continuous differential coefficient

$$\frac{\partial}{\partial t}\overline{n^k}(t, s, n') = \sum_{n=0}^{\infty}{}' n^k \left(n\left|\frac{\partial}{\partial t}\boldsymbol{P}(t, s)\right|n'\right). \tag{14}$$

In exactly the same way we can prove, starting only from the equation adjointed to the fundamental equation, (2.1.2a), that

$$\frac{\partial}{\partial s}\overline{n^k}(t, s, n') = \sum_{n=0}^{\infty}{}' n^k \left(n\left|\frac{\partial}{\partial s}\boldsymbol{P}(t, s)\right|n'\right). \tag{15}$$

We have thus proved

Theorem 2.

The moments (1) *of a generalized elementary stochastic process satisfying* (7) *not only exist and are continuous in each interval satisfying* (8), *but are, furthermore, differentiable with continuous differential coefficients.*

By the practical computation of the moments the following theorem is often useful (cf. chap. 4).

Theorem 3.

The moments (1) *of a generalized elementary stochastic process satisfying* (7) *not only exist, are continuous and differentiable with continuous differential coefficients in each interval satisfying* (8), *but, furthermore, satisfy the differential equations*

$$\frac{\partial}{\partial t}\overline{n^k}(t, s, n') = \sum_{n=0}^{\infty} n^k\left(n|\frac{\partial}{\partial t}\boldsymbol{P}(t, s)|n'\right) =$$

$$\sum_{n=0}^{\infty}\sum_{n''=0}^{\infty} n^k\big(n|\boldsymbol{A}(t)|n''\big)\big(n''|\boldsymbol{P}(t, s)|n'\big) =$$

$$\sum_{n''=0}^{\infty}\left(\sum_{n=0}^{n''+l} n^k\big(n|\boldsymbol{A}(t)|n''\big)\right)\big(n''|\boldsymbol{P}(t, s)|n'\big) \tag{16}$$

for all $n' = 0, 1, 2, \ldots$
,, ,, $t \geq s$
,, ,, $k \geq 0$.

Due to (14) we only have to prove that the inversion of the order of summation on the right hand side of (16) is legitimate. This is in fact the case on the assumption (7) because the double series in (16) will be absolutely convergent. We have, namely, due to (2.9.14), (2.1.6) and (7),

$$\sum_{n=0}^{\infty}\sum_{n''=0}^{\infty} n^k\big(n\big||\boldsymbol{A}(t)|\big|n''\big)\big(n''|\boldsymbol{P}(t, s)|n'\big) =$$

$$\sum_{n''=0}^{\infty}\left(\sum_{n=0}^{n''+l} n^k\big(n|(\boldsymbol{\Pi}+\boldsymbol{1})\cdot\boldsymbol{p}|n''\big)\right)(n''|\boldsymbol{P}|n') \leq \sum_{n''=0}^{\infty}(n''+l)^k 2p(n'', t)(n''|\boldsymbol{P}|n') \leq$$

$$2^{k+1}f(t)l^k\sum_{n''=1}^{\infty} n''^{k+1}(n''|\boldsymbol{P}|n') + l^k 2p(0, t)(0|\boldsymbol{P}|n') \tag{17}$$

for all $n' = 0, 1, 2, \ldots$
,, ,, $t \geq s$
,, ,, $k \geq 0$,

because

$$(n''+l)^k = \sum_{i=0}^{k} \binom{k}{i} n''^{k-i} l^i \leq n''^k l^k \sum_{i=0}^{k} \binom{k}{i} = n''^k l^k 2^k . \quad (18)$$

Due to theorem 1 all the moments exist and the series (17) is, therefore, convergent for all values of k, q. e. d.

CHAPTER 3.

The general Theory of multi-dimensional Stochastic Processes.

§ 3.1. In chap. 2 we have given the theory of one-dimensional stochastic processes, i.e. processes in which only one stochastic variable enters. We shall now in this chapter generalize our one-dimensional theory to a theory of multi-dimensional stochastic processes of the discontinuous type in which each of the stochastic variables can assume only an enumerable manifold of values. In a biological picture such processes can, as mentioned in the introduction, be interpreted in the following way. We consider a population consisting of individuals belonging to *different* species. Each individual can then disappear, be created or propagate and by these processes there may at the same time disappear or be created one or more individuals of the same or of the other species. Our stochastic variables are then accounting for the numbers of individuals of the different species present at the time t.

Let the number of the stochastic variables entering be denoted by m $(= 1, 2, 3, \ldots)$. As we have seen in chap. 1 such processes are again described by two infinite systems of simultaneous linear differential equations of the first order, being generalizations of (2.1.1) and (2.1.2)

$$\frac{\partial}{\partial t} P((n), t; (n'), s) = -p((n), t)P((n), t; (n'), s) + \sum_{n_1''=0}^{\infty} \ldots \sum_{n_m''=0}^{\infty} \Pi((n); (n''), t)p((n''), t)P((n''), t; (n'), s) \tag{1}$$
$$(n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots \; ; \; t \geq s)$$

and

$$\frac{\partial}{\partial s} P((n), t; (n'), s) = P((n), t; (n'), s)p((n'), s) - \sum_{n_1''=0}^{\infty} \ldots \sum_{n_m''=0}^{\infty} P((n), t; (n''), s)\Pi((n''); (n'), s)p((n'), s) \tag{2}$$
$$(n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots \; ; \; t \geq s)$$

in which we shall call (2) the equation **adjointed** to (1).

The symbols entering into (1) and (2) have the following meaning:

$P(n_1, \ldots, n_m, t; n_1', \ldots, n_m', s) = P((n), t; (n'), s)$ *is the conditioned, or relative, probability of the m stochastic variables assuming the values* $n_1, \ldots, n_m$, *respectively, at the time t relative to the hypothesis that they assume the values* $n_1', \ldots, n_m'$, *respectively, at the time s.*

$p(n_1, \ldots, n_m, t)\Delta t = p((n), t)\Delta t$ *is an asymptotic expression of the probability of a stochastic change of the variables taking place in the interval between t and* $t+\Delta t$ *when the variables assume the values* $n_1, \ldots, n_m$, *respectively, at the time t;* $p((n), t)$ *is called the* **intensity function.**

$\Pi(n_1, \ldots, n_m; n_1', \ldots, n_m', t) = \Pi((n); (n'), t)$ *is the conditioned probability of the variables assuming the values* $n_1, \ldots, n_m$, *respectively, at the time* $t+\Delta t$ *relative to the hypothesis that a stochastic change of the variables from the state* $n_1', \ldots, n_m'$ *has taken place during the interval between t and* $t+\Delta t$. *We shall call* $\Pi((n); (n'), t)$ *the* **relative transition probability from** *the state* $(n_1', \ldots, n_m') = (n')$ **to** *the state* $(n_1, \ldots, n_m) = (n)$.

By means of these definitions the equations (1) and (2) have, just as in the case of the one-dimensional theory, an immediate 'anschaulich' interpretation, representing 'continuity equations' in the positive and negative time direction, respectively, for the 'probability mass', expressing that this 'mass' can neither be created nor disappear. On the left hand side of (1) we have the rate of change with respect to t and this must equal the 'probability current' which 'flows into' the state (n) from all the other states, i. e. second term on right, minus the 'probability current' which 'flows out' from the state (n) to all the other states, i. e. first term on right. In the same way the left hand side of (2) represents the rate of change with respect to s and this must equal the 'probability current'

which 'flows out' from the state (n') to all the other states, i. e. first term on right, minus the 'probability current' which 'flows into' the state (n') from all the other states, i. e. second term on right.

From the definition of the p and Π functions it follows that

$$p((n), t) \geqq 0 \quad \text{for all } n_1, \ldots, n_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s, \tag{3}$$

$$0 \leqq \Pi((n); (n'), t) \leqq 1 \text{ for all } n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s, \tag{4}$$

$$\Pi((n'); (n'), t) = 0 \quad \text{for all } n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s \tag{5}$$

and

$$\sum_{n_1=0}^{\infty} \ldots \sum_{n_m=0}^{\infty} \Pi((n); (n'), t) = 1 \quad \text{for all } n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s. \tag{6}$$

These four relations are direct generalizations of (2.1.3)-(2.1.6).

It is now, just as in the case of the one-dimensional theory, the object of our theory to show that if all the functions $p((n), t)$ and $\Pi((n); (n'), t)$ are given, satisfy (3)–(6) and are assumed to be continuous for all $t \geqq s$, then (1) and (2) each has one and only one solution which is the same for both systems and which satisfies the following conditions being necessary and sufficient for the solutions denoting a probability distribution for a discontinuous stochastic process

$$\lim_{t \to s} P((n), t; (n'), s) = \lim_{t \leftarrow s} P((n), t; (n'), s) = \delta_{n_1, n_1'} \cdots \delta_{n_m, n_m'} \tag{I}$$
$$\text{for all } n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots$$

$$P((n), t+\Delta t; (n'), t) = \begin{cases} 1 - p((n'), t)\Delta t + o(\Delta t) & \text{for } (n_1, \ldots, n_m) = (n_1', \ldots, n_m') \\ \Pi((n); (n'), t)p((n'), t)\Delta t + o(\Delta t) & \text{for } (n_1, \ldots, n_m) \neq (n_1', \ldots, n_m') \end{cases} \tag{II}$$

$$P((n), t; (n'), s) = \sum_{n_1''=0}^{\infty} \ldots \sum_{n_m''=0}^{\infty} P((n), t; (n''), \tau) P((n''), \tau; (n'), s) \tag{III}$$
$$\text{for all } n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \tau \text{ in } s \leqq \tau \leqq t$$

(Chapman-Kolmogoroff's equation)

$$0 \leqq P((n), t; (n'), s) \leqq 1 \text{ for all } n_1, \ldots, n_m, n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s \tag{IV}$$

$$\sum_{n_1=0}^{\infty} \ldots \sum_{n_m=0}^{\infty} P((n), t; (n'), s) = 1 \quad \text{for all } n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad ,, \quad ,, \quad t \geqq s. \tag{V}$$

These relations are direct generalizations of (2.1.I)–(2.1.V) and may, just like the latter, be written in a far more compact form using the *matrix symbolism*[1]). We first introduce the m-dimensional **distribution matrix**

$$\boldsymbol{P}(t, s) = \{((n)|\boldsymbol{P}(t, s)|(n'))\} = \{P((n), t; (n'), s)\}. \qquad (7)$$

Next, we introduce the m-dimensional **intensity matrix** as the m-dimensional diagonal matrix

$$\boldsymbol{p}(t) = \{((n)|\boldsymbol{p}(t)|(n'))\} = \{p((n'), t)\delta_{n_1, n_1'} \cdots \delta_{n_m, n_m'}\} \qquad (8)$$

and finally the m-dimensional **relative transition matrix**

$$\boldsymbol{\Pi}(t) = \{((n)|\boldsymbol{\Pi}(t)|(n'))\} = \{\Pi((n); (n'), t)\}. \qquad (9)$$

As generalizations of the matrix operations defined in (2.1.10) we now define the **column sum** and the **row sum matrices** by

$$\boldsymbol{\Sigma} \cdot \boldsymbol{M} = \left\{ \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} ((n)|\boldsymbol{M}|(n')) \right\} = \{(\boldsymbol{\Sigma} \cdot \boldsymbol{M}|(n'))\}$$

and (10)

$$\boldsymbol{M} \cdot \boldsymbol{\Sigma} = \left\{ \sum_{n_1'=0}^{\infty} \cdots \sum_{n_m'=0}^{\infty} ((n)|\boldsymbol{M}|(n')) \right\} = \{((n)|\boldsymbol{M} \cdot \boldsymbol{\Sigma})\}.$$

We can now write our **fundamental equations** (1) and (2) in the form

$$\frac{\partial}{\partial t} \boldsymbol{P}(t, s) = -\boldsymbol{p}(t) \cdot \boldsymbol{P}(t, s) + \boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t) \cdot \boldsymbol{P}(t, s) = \boldsymbol{A}(t) \cdot \boldsymbol{P}(t, s) \qquad (1a)$$

and

$$\frac{\partial}{\partial s} \boldsymbol{P}(t, s) = \boldsymbol{P}(t, s) \cdot \boldsymbol{p}(s) - \boldsymbol{P}(t, s) \cdot \boldsymbol{\Pi}(s) \cdot \boldsymbol{p}(s) = -\boldsymbol{P}(t, s) \cdot \boldsymbol{A}(s) \qquad (2a)$$

or

$$\frac{\partial}{\partial s} \boldsymbol{P}^*(t, s) = -\boldsymbol{A}^*(s) \cdot \boldsymbol{P}^*(t, s). \qquad (2b)$$

(2b) shows that *equation* (2) *is of exactly the same form as* (1).

[1]) For the definition and computation rules of m-dimensional matrices we refer to § 7.5. As for one-dimensional matrices the former will themselves be denoted by clarendon letters, $\boldsymbol{P}, \boldsymbol{p}, \boldsymbol{\Pi}$ etc. and their elements by the DIRAC symbols $((n)|\boldsymbol{P}|(n'))$, and so on. $n_1, \ldots, n_m = (n)$ are called the **row indices** and can be interpreted as a point in an m-dimensional n-space, and $n_1', \ldots, n_m' = (n')$ the **column indices** and can be interpreted as a point in another m-dimensional n'-space. We note that the **m-dimensional unit-matrix 1** is, of course, defined by $((n)|\boldsymbol{1}|(n')) = \delta_{n_1, n_1'} \cdots \delta_{n_m, n_m'}$. Furthermore, we shall write $\boldsymbol{P} = \{((n)|\boldsymbol{P}|(n'))\}$ etc.

The m-dimensional **operator matrix** introduced in (1a) and (2a)

$$\boldsymbol{A}(t) = -\boldsymbol{p}(t) + \boldsymbol{\Pi}(t) \cdot \boldsymbol{p}(t) = (\boldsymbol{\Pi}(t) - \boldsymbol{1}) \cdot \boldsymbol{p}(t) \tag{11}$$

has the elements

$$\big((n)|\boldsymbol{A}(t)|(n')\big) = \begin{cases} \Pi\big((n); (n'), t\big) p\big((n'), t\big) \geqq 0 \text{ for } (n_1, \ldots, n_m) \neq (n_1', \ldots, n_m') \\ -p\big((n'), t\big) \leqq 0 \text{ for } (n_1, \ldots, n_m) = (n_1', \ldots, n_m'). \end{cases} \tag{12}$$

Furthermore, due to (6) it satisfies

$$\boldsymbol{\Sigma} \cdot \boldsymbol{A}(t) = \left\{ \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} \big((n)|\boldsymbol{A}(t)|(n')\big) \right\} = \boldsymbol{0} \tag{13}$$

$$\text{for all } n_1', \ldots, n_m' = 0, 1, 2, \ldots .$$

The **fundamental conditions** (I)–(V) can now be written

$$\lim_{t \to s} \boldsymbol{P}(t, s) = \lim_{t \leftarrow s} \boldsymbol{P}(t, s) = \boldsymbol{1} \tag{Ia}$$

$$\boldsymbol{P}(t + \Delta t, t) = \boldsymbol{1} + \boldsymbol{A}(t)\Delta t + \boldsymbol{o}(\Delta t) \quad \text{for all } t \geqq s \tag{IIa}$$

$$\boldsymbol{P}(t, s) = \boldsymbol{P}(t, \tau) \cdot \boldsymbol{P}(\tau, s) \quad \text{for all } s \leqq \tau \leqq t \tag{IIIa}$$

$$\boldsymbol{0} \leqq \boldsymbol{P}(t, s) \leqq \{1\} \quad \text{for all } t \geqq s \tag{IVa}$$

$$\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s) = \{1\} \quad \text{for all } t \geqq s. \tag{Va}$$

We now see that having introduced the matrix symbolism, the *dimension has disappeared from our equations*, the equations (1a)–(2b), (11), (13) and (Ia)–(Va) being *identical* with the corresponding ones in chap. 2. This fact indicates that the theory of multi-dimensional processes is contained in the theory of one-dimensional processes. This is actually the case due to the fact that the *conception of dimension is an artificial conception from the point of view of the theory of manifolds* as shown by the fundamental theorem that

$$\text{(enumerable infinite)}^m = \text{enumerable infinite}. \tag{14}$$

Due to (14) *we can simply re-number both the row and the column indices* 0, 1, 2, ... *and thus transform our equations from the m-dimensional to a one-dimensional form*. We emphasize that this re-numbering shall, of course, be the same for the row index as for the column index space. Otherwise, corresponding factors in the products would not have the same partners after the re-numbering.

We now assume, as in chap. 2, *that the m-dimensional operator matrix* $\boldsymbol{A}(t)$ *given in* (11) *is* **absolutely exponentiable** *in some interval* $s \leq \tau \leq t$ which means (cf. § 7.5) that the matrix function

$$\exp[\boldsymbol{K}(t-s)] = \sum_{\nu=0}^{\infty} \boldsymbol{K}^{\nu} \frac{(t-s)^{\nu}}{\nu!} \tag{15}$$

exists, when

$$\boldsymbol{K} = \max_{s \leq \tau \leq t} |\boldsymbol{A}(\tau)|. \tag{16}$$

Because of the fundamental property of absolutely convergent manifold series that the order of summation is irrelevant, it is then obvious that *the exponentiability is invariant against any such renumbering*. Consequently, the theory of m-dimensional stochastic processes is contained in the theory of one-dimensional processes. In fact, we now see that all the contents of §§ 2.2–2.6—except the last part, eqs. (2.6.12)–(2.6.21)—may just as well be interpreted as applying to m-dimensional as to one-dimensional matrices, as the dimension does not enter explicitly. From § 2.7 we may at once take over the following theorem.

Theorem 1.

For an arbitrary m-dimensional stochastic process, the operator matrix of which satisfies (15), *i. e. is absolutely exponentiable, the fundamental equation* (Ia) *has one and only one solution, given by the product-integral*

$$\boldsymbol{P}(t, s) = \mathcal{P}_{s}^{t}(\boldsymbol{1} + \boldsymbol{A}(\tau)d\tau) = \lim_{m' \to \infty} \prod_{i=0}^{m-1} (\boldsymbol{1} + \boldsymbol{A}(\tau_i) l_i). \tag{17}$$

This solution, being also the unique solution of (2a), *satisfies the fundamental conditions* (Ia)–(IVa) *and the condition*

$$(\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s) | (n')) \leq 1 \qquad \text{for all } n_1', \ldots, n_m' = 0, 1, 2, \ldots \quad \text{,, ,, } t \geq s\,. \tag{18}$$

The amount missing in (18) we can again interpret as the probability of the value (∞)

$$P((\infty), t; (n'), s) = 1 - (\boldsymbol{\Sigma} \cdot \boldsymbol{P}(t, s) | (n')) \geq 0. \tag{19}$$

We observe that the notions of a *finite*, of a *pure absorption* and of a *pure propagation* process may also be taken over into the m-dimensional theory. By the first process we understand, of course, a process the operator matrix of which is an m-dimensional *finite*

matrix of some order N (cf. (7.5.13) and (7.5.14)). By the second process we understand an m-dimensional process in which *neither* of the m stochastic variables increases whenever a change takes place. From (7.5.26) we see that such processes are characterized by having m-dimensional *column half* operator matrices. In the same way *neither* of the m stochastic variables *decreases* in the third process whenever a change takes place. The operator matrix of the pure propagation process is thus, cf. (7.5.25), an m-dimensional *row half* matrix. For the finite and for the pure absorption processes all summations in (18) are finite and the calculation (2.8.2)–(2.8.3) may obviously in this case be generalized immediately. We thus have

Theorem 2.

For finite and for pure absorption processes the sign of equality holds good in (18), *i. e. the condition* (Va) *is satisfied.*

§ 3.2. We shall now generalize the FELLER-LUNDBERG condition for the validity of (3.1.Va) (cf. § 2.9). In the m-dimensional case the notion of an *elementary* process is, so far as we can see, of no special interest.

It could, of course, easily be transferred to denote an m-dimensional pure propagation process in which each variable changes by $+1$ whenever a change takes place. The operator matrix would thus be both a row half (cf. (7.5.25)) and a column semi-diagonal matrix with $l = 1$ (cf. (7.5.20)).

The notion of a **generalized elementary *m*-dimensional stochastic process** is, on the other hand, an important notion. By this term we shall, of course, understand an m-dimensional stochastic process in which neither of the m variables can change, when jumping *upwards*, by more than a certain number of stages, say l, common for all the variables and independent of the state before the change. $\Pi(t)$ and, consequently also $A(t)$, is thus assumed to be an m-dimensional **column semi-diagonal** matrix (cf. (7.5.20)). We shall, as in the one-dimensional theory, confine ourselves to the consideration of such generalized elementary stochastic processes as this case will–presumably–cover all cases that may be met with in the practical applications of the theory. As a generalization of (2.9.8) we now have

$$0 \leq 1 - \sum_{n_1=0}^{N_1} \cdots \sum_{n_m=0}^{N_m} \big((n)|\boldsymbol{P}(t, s)|(n')\big) =$$

$$\left(\sum_{n_1=N_1+1}^{\infty} \cdots \sum_{n_m=N_m+1}^{\infty} \; \sum_{n''_1=0}^{N_1} \cdots \sum_{n''_m=0}^{N_m} - \sum_{n_1=0}^{N_1} \cdots \sum_{n_m=0}^{N_m} \; \sum_{n''_1=N_1+1}^{\infty} \cdots \sum_{n''_m=N_m+1}^{\infty} \right) \cdot$$

$$\int_s^t \big((n)|\boldsymbol{\Pi}(\tau) \cdot \boldsymbol{p}(\tau)|(n'')\big)\big((n'')|\boldsymbol{P}(\tau, s)|(n')\big)d\tau =$$

$$\text{'up'—'down'} < 1 \qquad \begin{array}{l} \text{for all } n'_1, \ldots, n'_m = 0, 1, 2, \ldots \\ \;\;,, \quad ,, \;\; N_i \geq n'_i, \; i = 1, 2, \ldots, m \\ \;\;,, \quad ,, \;\; t \geq s \; . \end{array} \tag{1}$$

LUNDBERG's argument can now immediately be generalized. In fact we obtain from (1), using (3.1.6), as a generalization of (2.9.15)

$$0 \leq \delta = 1 - \big(\Sigma \cdot \boldsymbol{P}(t, s)|(n')\big) = P\big((\infty), t; (n'), s\big) \leq$$

$$\sum_{n''_1=N_1-l+1}^{N_1} \cdots \sum_{n''_m=N_m-l+1}^{N_m} \int_s^t p\big((n''), \tau\big)\big((n'')|\boldsymbol{P}(\tau, s)|(n')\big)d\tau \leq$$

$$\bar{p}\big((N)\big) \sum_{n''_1=N_1-l+1}^{N_1} \cdots \sum_{n''_1=N_m-l+1}^{N_m} \int_s^t \big((n'')|\boldsymbol{P}(\tau, s)|(n')\big)d\tau \; . \tag{2}$$

$$\begin{array}{l} \text{for all } n'_1, \ldots, n'_m = 0, 1, 2, \ldots \\ \;\;,, \quad ,, \;\; N_i \geq n'_i, \; i = 1, 2, \ldots, m \\ \;\;,, \quad ,, \;\; t \geq s \; . \end{array}$$

In (2) we have now defined, instead of (2.9.16),

$$\bar{p}\big((n)\big) = \max_{s \leq \tau \leq t} \big\{p\big((n'), \tau\big)\big\} \text{ for} \left\{ \begin{array}{l} n'_1 = n_1 - l + 1, \; n_1 - l + 2, \; \ldots, \; n_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ n'_m = n_m - l + 1, \; n_m - l + 2, \; \ldots, \; n_m \end{array} \right\} . \tag{3}$$

Dividing (2) with $\bar{p}$ and summing over all $N_1, \ldots, N_m$ we obtain, using (3.1.18), as a generalization of (2.9.17)

$$\delta \sum_{N_1=n'_1}^{\infty} \cdots \sum_{N_m=n'_m}^{\infty} \frac{1}{\bar{p}\big((N)\big)} \leq$$

$$\int_s^t \sum_{n''_1=0}^{l-1} \cdots \sum_{n''_m=0}^{l-1} \sum_{N_1=n'_1}^{\infty} \cdots \sum_{N_m=n'_m}^{\infty} \big((N-n'')|\boldsymbol{P}(\tau, s)|(n')\big)d\tau \leq ml(t-s) < \infty \tag{4}$$

$$\begin{array}{l} \text{for all } n'_1, \ldots, n'_m = 0, 1, 2, \ldots \\ \;\;,, \quad ,, \;\; t \geq s \; . \end{array}$$

(4) shows that we have proved

Theorem 1.

For m-dimensional generalized elementary stochastic processes, i. e. processes the operator matrix of which is an m-dimensional column semi-diagonal matrix, a **sufficient** *condition for the validity of the sign of equality in* (3.1.18), *i. e. the validity of* (3.1.Va), *in the whole interval* $s \leqq \tau \leqq t$ *consists in the divergence of the series*

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} \frac{1}{\bar{p}((n))} = \infty \tag{5}$$

in which $\bar{p}((n))$ *is given in* (3).

As corollaries of theorem 1 we note the following which ensue in exactly the same way as in the one-dimensional theory and which are often met with in practice.

Corollary 1.

Let for an m-dimensional generalized elementary stochastic process $\bar{p}((n))$ *given in* (3) *satisfy the condition*

$$\bar{p}((n)) \leqq \sum_{i=1}^{m} c_i n_i \quad \text{for all } n_j \geqq N_j, \quad j = 1, 2, \ldots, m \tag{6}$$

in which $c_1, \ldots, c_m$ *are arbitrary non-negative constants* $((c) \neq (0))$ *and* $N_1, \ldots, N_m$ *are arbitrary non-negative integers* $((N) \neq (0))$. *Then* (5) *is fulfilled and* (3.1.Va) *consequently valid.*

(6) shows that the operator matrix is of type IV (cf. § 7.6) and we thus see that (3.1.Va) is valid in this case.

Corollary 2.

Let for an m-dimensional generalized elementary stochastic process $\bar{p}((n))$ *given in* (3) *satisfy the condition*

$$\bar{p}((n)) \leqq c \quad \text{for all } n_1, \ldots, n_m = 0, 1, 2, \ldots \tag{7}$$

in which $c > 0$ *is an arbitrary constant. Then* (5) *is fulfilled and* (3.1.Va) *consequently valid.*

(7) shows that the operator matrix is *bounded*, i. e. of type I, and we thus see that (3.1.Va) is valid in this case.

Finally, we note that theorem 1 § 2.10 is, obviously, generalized to

Theorem 2.

For an **arbitrary,** *and not necessarily generalized elementary, m-dimensional pure propagation process, i. e. a process the operator*

matrix of which is an m-dimensional row half matrix, a **necessary** *condition for the validity of* (3.1.Va) *in the whole interval* $s \leq \tau \leq t$ *consists in the divergence of the series*

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} \frac{1}{\underline{p}((n))} = \infty \tag{8}$$

in which

$$\underline{p}((n)) = \min_{s \leq \tau \leq t} p((n), \tau). \tag{9}$$

We leave the proof together with the generalization of theorem 2 § 2.10 to the reader.

§ 3.3. Finally, we shall discuss the question of the moments of an m-dimensional probability distribution. The **moments of order** $\boldsymbol{k}$ of the relative distribution $\boldsymbol{P}(t, s)$ are defined by the following generalizations of (2.11.1)

$$\overline{n_1^{k_1} \cdots n_m^{k_m}}(t, s, (n')) = \sum_{n_1=0}^{\infty} \cdots \sum_{n_m=0}^{\infty} n_1^{k_1} \cdots n_m^{k_m} ((n)|\boldsymbol{P}(t, s)|(n')) \geq 0 \tag{1}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$
" " $t \geq s$
" " non-negative k_i satisfying $k_1 + \ldots + k_m = k$
" " $k \geq 0$.

Corresponding to (2.11.4) we have from (3.1.18)

$$\overline{n_1^0 \cdots n_m^0}(t, s, (n')) = (\Sigma \cdot \boldsymbol{P}(t, s)|(n')) \leq 1 \tag{2}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$.

Corresponding to (2.11.6) we have from (3.1.Ia)

$$\overline{n_1^{k_1} \cdots n_m^{k_m}}(s, s, (n')) = n_1'^{k_1} \cdots n_m'^{k_m} \tag{3}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$.

As a generalization of theorem 1 § 2.11 we now have the following theorem which is a natural extension of theorem 4 § 7.5.

Theorem 1.

For an m-dimensional, generalized elementary stochastic process whose intensity function satisfies

$$p((n), t) \geq f(t) \cdot \sum_{i=1}^{m} c_i n_i \quad \text{for all } (n_1, \ldots, n_m) \neq (0, \ldots, 0), \tag{4}$$

i. e. whose operator matrix is of type IV *(cf.* § 7.6*), we have for all* $k \geq 0$ *that all the moments exist and are continuous in each interval* (s, t) *for which*

$$t - s < \frac{1}{\left(\sum_{i=1}^{m}{}' c_i\right) lC} \tag{5}$$

with

$$C = 2 \max_{s \le \tau \le t} f(\tau) . \tag{6}$$

For, instead of (2.11.10) we have, using (7.5.22) and lemma 4 § 7.5,

$$\overline{n_1^{k_1} \cdots n_m^{k_m}}(\tau, s, (n')) \le \sum_{n_1=0}^{\infty}{}' \cdots \sum_{n_m=0}^{\infty}{}' n_1^{k_1} \cdots n_m^{k_m}((n)|\mathbf{exp}[\boldsymbol{K}(t-s)]|(n')) =$$

$$\sum_{\nu=0}^{\infty}{}' \left(\sum_{n_1=0}^{n_1'+\nu l}{}' \cdots \sum_{n_m=0}^{n_m'+\nu l}{}' n_1^{k_1} \cdots n_m^{k_m}((n)|\boldsymbol{K}^{\nu}|(n')) \right) \frac{(t-s)^{\nu}}{\nu!} \le$$

$$\sum_{\nu=0}^{\infty}{}' (n_1' + \nu l)^{k_1} \cdots (n_m' + \nu l)^{k_m} M_{(n')} \prod_{\beta=1}^{\nu-1} \left(\sum_{i=1}^{m}{}' c_i (n_i' + \beta l) \right) C^{\nu-1} \frac{(t-s)^{\nu}}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}{}' v_{\nu} \tag{7}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, τ in $s \le \tau \le t$
,, ,, non-negative k_i satisfying $k_1 + \ldots + k_m = k$
,, ,, $k \ge 0$.

This series is, of course, convergent in the interval (5) because we have, as a generalization of (2.11.11) and (7.5.41),

$$\frac{v_{\nu+1}}{v_{\nu}} = \left(\frac{n_1' + (\nu+1)l}{n_1' + \nu l}\right)^{k_1} \cdots \left(\frac{n_m' + (\nu+1)l}{n_m' + \nu l}\right)^{k_m} \cdot \frac{\sum_{i=1}^{m}{}' c_i (n_i' + \nu l)}{\nu + 1} C(t-s) \underset{\nu \to \infty}{\to}$$

$$\left(\sum_{i=1}^{m}{}' c_i\right) lC(t-s) < 1 \quad \text{for all } k \ge 0 . \tag{8}$$

Thus, all the moments exist, and as (7) shows that the convergence is uniform in $s \le \tau \le t$ they are, consequently, continuous, q. e. d.

(8) shows still more, namely, that the *radius of convergence* of the series (7) is *independent* of k, being the same for all the moments. We can thus use exactly the **same** *exponentiable continuation* as for $\boldsymbol{P}(t, s)$ itself (cf. § 2.7). We thus find that the restriction (5) is, in general, artificial, *the moments existing throughout the region of definition of the operator matrix*.

Finally, we give the m-dimensional generalizations of theorems 2 and 3 § 2.11.

Theorem 2.

The moments (1) *of an m-dimensional, generalized elementary stochastic process satisfying* (4) *not only exist, are continuous and differentiable with continuous differential coefficients in each interval satisfying* (5), *but, furthermore, satisfy the differential equations*

$$\frac{\partial}{\partial t}\overline{n_1^{k_1}\cdots n_m^{k_m}}(t,s,(n')) = \sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty} n_1^{k_1}\cdots n_m^{k_m}\Big((n)|\frac{\partial}{\partial t}\boldsymbol{P}(t,s)|(n')\Big) =$$

$$\sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty}\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty} n_1^{k_1}\cdots n_m^{k_m}((n)|\boldsymbol{A}(t)|(n''))((n'')|\boldsymbol{P}(t,s)|(n')) =$$

$$\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty}\left(\sum_{n_1=0}^{n_1''+l}\cdots\sum_{n_m=0}^{n_m''+l} n_1^{k_1}\cdots n_m^{k_m}((n)|\boldsymbol{A}(t)|(n''))\right)((n'')|\boldsymbol{P}(t,s)|(n')) \tag{9}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$
,, ,, $t \geqq s$
,, ,, non-negative k_i satisfying $k_1 + \ldots + k_m = k$
,, ,, $k \geqq 0$.

The proof of the first part of the theorem consists of an obvious generalization of (2.11.12). The second part is proved by the generalization of (2.11.17) which gives, using (2.11.18),

$$\sum_{n_1=0}^{\infty}\cdots\sum_{n_m=0}^{\infty}\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty} n_1^{k_1}\cdots n_m^{k_m}((n)|\,|\boldsymbol{A}|\,|(n''))((n'')|\boldsymbol{P}|(n')) =$$

$$\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty}\left(\sum_{n_1=0}^{n_1''+l}\cdots\sum_{n_m=0}^{n_m''+l} n_1^{k_1}\cdots n_m^{k_m}((n)|(\boldsymbol{\Pi}+\boldsymbol{1})\cdot\boldsymbol{p}|(n''))\right)((n'')|\boldsymbol{P}|(n')) \leqq$$

$$\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty}(n_1''+l)^{k_1}\cdots(n_m''+l)^{k_m}2p((n''),t)((n'')|\boldsymbol{P}|(n')) \leqq$$

$$2^{k+1}f(t)l^k\sum_{i=1}^{m}c_i\sum_{n_1''=0}^{\infty}\cdots\sum_{n_m''=0}^{\infty} n_1''^{k_1}\cdots n_m''^{k_m}n_i''((n'')|\boldsymbol{P}|(n'))+$$

$$l^k 2p((0),t)((0)|\boldsymbol{P}|(n')) \tag{10}$$

for all $n_1', \ldots, n_m' = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, $t \geqq s$
,, ,, non-negative k_i satisfying $k_1 + \ldots + k_m = k$
,, ,, $k \geqq 0$.

Due to theorem 1 all the moments exist and the series (10) is, therefore, convergent for all values of k. The inversion of the order of summation in (9) is thus legitimate, q. e. d.

PART II

PHYSICAL THEORY

CHAPTER 4.

Discussion of special Stochastic Processes used in the Theory of Cosmic Radiation.

§ 4.1. We shall in this chapter discuss some special stochastic processes which we shall later on apply to the theory of that part of the cosmic radiation called the *soft component*. This component is now known to consist of high-speed *electrons*[1]) and very energetic electromagnetic radiation quanta called *photons*. As will be discussed in the next chapter an energetic photon has by travelling a distance Δt in some material substance a certain chance of being absorbed by emitting a pair consisting of one positron and one negatron. The probability of this process is asymptotically proportional to Δt, the factor of proportionality being a function of the energies of the 'parent' photon and the resulting electron 'off-spring' (cf. chap.5). An energetic electron has by travelling a distance Δt in the same material substance a certain chance of losing some energy by emitting one photon. The probability of this process is again asymptotically proportional to Δt and the factor of proportionality is again a function of the respective energies (cf. chap. 5). If now a primary energetic electron (or photon) enters into a layer of the material substance, it will be seen that the result of the combination of these elementary processes will be the formation of a whole 'family' of secondary electrons and photons, a socalled *cascade* or *multiplication shower*. The primary electron has, namely, a certain chance of emitting a photon, which has a certain chance of being absorbed by emitting a pair of electrons, each of which again has a certain chance of emitting photons and so on, as illustrated on fig. 2.

It will be seen that this cascade process is to some extent analogous

[1]) By 'electrons' we shall understand both positive and negative electrons. The former we shall call 'positrons', the latter 'negatrons'.

to biological processes in which one generation in the course of time gives way to another. The time in the biological analogy will thus correspond to the thickness of the material layer, for to a first approximation the direction of motion is known to be conserved
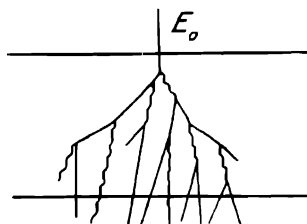
FIG. 2. *Schematic picture of a multiplication shower initiated by a primary electron with energy E_0. (Straight lines: electrons. Wavy lines: photons.)*

by each of the elementary cascade processes. What is measured experimentally is, however, only the electrons and we, consequently, here have to do with a feature also met with in biological processes (under the name of *heterogony*), namely that the generations alternate, the members of two successive generations being of different kinds belonging, so to speak, to different species.

As shown in the next chapter it is possible to evaluate the *average numbers* of electrons and photons, with given energies, which are produced from a parent electron (or photon), with definite energy, in a certain layer of material. This calculation is, however, already so complicated that it would be quite hopeless to try to evaluate directly the probabilities of finding given numbers of electrons and photons as functions of the primary and secondary energies and of the thickness of the layer. The problem is, therefore, to construct a *simplified model of the multiplication process* which on the one hand retains most of its characteristic features and on the other hand permits of numerical results being obtained.

§ 4.2. The first simplified model was discussed by BHABHA and HEITLER in their fundamental paper on the theory of cosmic ray showers[1]). They assumed that in the shower the probability for a secondary electron to emerge with energy between E_1 and $E_1 + \Delta E_1$ is independent of the probability for another secondary electron to emerge with energy between E_2 and $E_2 + \Delta E_2$. It may, of course,

[1]) BHABHA and HEITLER (1937), in the following quoted as B & H.

be strongly doubted whether this assumption is really true since both electrons are generated in one multiplication process from the same initial parent electron and, therefore, the two probabilities mentioned are in fact in some way correlated. From this assumption of independence it follows at once, however, that the probability $P(n, t)$ of finding n *secondaries* at the depth t is given by the **Poisson distribution**

$$P(n, t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!} \qquad n = 0, 1, 2, \ldots \quad (\lambda t \geqq 0). \tag{1}$$

Here

$$\sum_{n=0}^{\infty} P(n, t) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} = e^{-\lambda t} e^{\lambda t} = 1, \tag{2}$$

and next the average number $\bar{n}$ is given by

$$\bar{n}(t) = \sum_{n=0}^{\infty} n P(n, t) = e^{-\lambda t} \lambda t \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} = \lambda t e^{-\lambda t} e^{\lambda t} = \lambda t. \tag{3}$$

For the second moment one finds

$$\overline{n^2}(t) = \sum_{n=0}^{\infty} n^2 P(n, t) = e^{-\lambda t} \lambda t \sum_{n=1}^{\infty} n \frac{(\lambda t)^{n-1}}{(n-1)!} = e^{-\lambda t} \lambda t \sum_{n=1}^{\infty} [(n-1)+1] \frac{(\lambda t)^{n-1}}{(n-1)!}$$

$$e^{-\lambda t} \lambda t \left[\sum_{n=2}^{\infty} \frac{(\lambda t)^{n-1}}{(n-2)!} + \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \right] = (\lambda t)^2 + \lambda t = \bar{n}^2 + \bar{n}, \tag{4}$$

and thus the fluctuation σ is given by

$$\sigma^2(t) = \overline{n^2} - \bar{n}^2 = \lambda t = \bar{n} \tag{5}$$

and the relative fluctuation δ by

$$\delta = \frac{\sigma}{\bar{n}} = \frac{1}{\sqrt{\lambda t}} = \frac{1}{\sqrt{\bar{n}}}. \tag{6}$$

The stochastic process corresponding to the distribution (1) is given by the following assumptions: (*a*) *the probability of a secondary electron being created in the small distance* Δt *is asymptotically proportional to* Δt, $\lambda \Delta t$, and (*b*) *this probability is independent of the number already present.* The mean value $\bar{n}$ given in (3) can now be deduced very easily from these assumptions. We have, namely, asymptotically that $\bar{n}(t+\Delta t)$ is a sum of two terms, $\bar{n}(t)$ and the amount created in the time interval Δt which on our assumptions

is equal to $\lambda \Delta t$[1]). We thus have the *continuity equation*

$$\bar{n}(t+\Delta t) = \bar{n}(t)+\lambda\Delta t+o(\Delta t)\,,$$
$$\text{i. e.}\quad \frac{d\bar{n}}{dt} = \lambda \quad \text{whence} \quad \bar{n} = \lambda t\,. \tag{7}$$

From the assumptions (a) and (b) it follows by means of the elementary rules for combining probabilities that $P(n, t+\Delta t)$ is equal to the product of the probabilities for n secondaries being present at t and 0 being created in Δt added to the product of the probabilities for $n-1$ being present at t and 1 being created in Δt. So we have

$$P(0,\, t+\Delta t) = P(0,\, t)(1-\lambda\Delta t)+o(\Delta t)$$
$$P(n,\, t+\Delta t) = P(n,\, t)(1-\lambda\Delta t)+P(n-1,\, t)\lambda\Delta t+o(\Delta t) \qquad n = 1, 2, 3, \ldots \tag{8}$$

which system for $\Delta t \to 0$ gives the system of simultaneous differential equations[2])

$$\frac{dP(0,t)}{dt} = -\lambda P(0,\, t)$$
$$\frac{dP(n,\, t)}{dt} = \lambda P(n-1,\, t)-\lambda P(n,\, t) \qquad n = 1, 2, 3, \ldots \quad (\lambda \geq 0,\; t \geq 0) \tag{9}$$

with the initial condition

$$P(n,\, 0) = \delta_{n,0} = \begin{cases} 1 & \text{for } n = 0 \\ 0 & \;\;,, \quad n \neq 0\,. \end{cases} \tag{10}$$

This system of equations governs the stochastic process which we shall call the **Poisson process.** From the general theory given in chap. 2 this system (whose operator matrix is of type I cf. § 7.6) is known to have a unique solution for which $0 \leq P(n, t) \leq 1$ and $\sum_{n=0}^{\infty} P(n, t) = 1$. In the present case the system of simultaneous differential equations (9) is, however, so simple that these facts can be shown directly. They are, namely, ordinary inhomogeneous linear

[1]) In Δt only one individual or none can be created, but in such case the mean value is equal to the probability itself, cf. e. g. A & B p. 39.

[2]) Comparing with the general theory in chap. 2 we see that this is a stochastic process with the intensity function $p(n, t) = \lambda$ and the relative transition probability $(n\,\Pi\,n') = \delta_{n',\,n+1}$.

differential equations of the first order and can, therefore, be integrated immediately to[1])

$$
\begin{aligned}
P(0,t) &= c_0 e^{-\lambda t} \\
P(n,t) &= e^{-\lambda t}\left(\lambda \int_0^t P(n-1,t')e^{\lambda t'}dt' + c_n\right) \qquad n = 1, 2, 3, \ldots .
\end{aligned} \tag{11}
$$

From (10) it follows that $c_0 = 1$, $c_n = 0$ and by induction it is seen at once that the solution of (11) just yields the expressions given in (1).

We note that in the symbolism used in chap. 2 the equations (9) of the Poisson process are written

$$
\frac{\partial}{\partial t}\boldsymbol{P}(t,s) = \boldsymbol{A}(t)\cdot\boldsymbol{P}(t,s) = \lambda \begin{Bmatrix} -1 & 0 & 0 & 0 & \ldots \\ 1 & -1 & 0 & 0 & \ldots \\ 0 & 1 & -1 & 0 & \ldots \\ 0 & 0 & 1 & -1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \cdot \boldsymbol{P}(t,s). \tag{12}
$$

As the operator matrix $\boldsymbol{A}(t)$ is here *constant* the solution is simply given by (2.6.7), i. e. by

$$
\boldsymbol{P}(t,s) = \exp\left[\boldsymbol{A}(t-s)\right] = \sum_{\nu=0}^{\infty} \boldsymbol{A}^\nu \frac{(t-s)^\nu}{\nu!}. \tag{13}
$$

It may be interesting to observe that this formula may be evaluated direct. We shall here carry out this calculation as an example of the evaluation of the product-integral. From (12) we have

$$
\boldsymbol{A} = \lambda(\boldsymbol{\theta} - \boldsymbol{1}) \tag{14}
$$

in which

$$
\boldsymbol{\theta} = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 1 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{Bmatrix}, \quad
\boldsymbol{\theta}^2 = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 1 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{Bmatrix}, \quad
\boldsymbol{\theta}^3 = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 0 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{Bmatrix} \text{ etc.} \tag{15}
$$

It is seen from (15) that when the power is increased by one the diagonal containing only ones is displaced one step *down*wards, i. e.

$$
(p|\boldsymbol{\theta}^\nu|q) = \delta_{p,q+\nu}. \tag{16}
$$

Because $\boldsymbol{\theta}$ and $\boldsymbol{1}$ commute we now have

$$
\boldsymbol{A}^\nu = \lambda^\nu \sum_{\mu=0}^{\nu} \binom{\nu}{\mu} \boldsymbol{\theta}^\mu (-\boldsymbol{1})^{\nu-\mu} \tag{17}
$$

[1]) Comparing with the general theory in chap. 2 we see that the Poisson process is a *pure propagation process* (cf. § 2.8) characterized by the operator matrix of the equations being a *row half matrix* (cf. § 7.3). The solution is then given in (2.6.21).

and inserting (16) and (17) into (13) we find

$$(n|\boldsymbol{P}(t,s)|n') =$$

$$\sum_{\nu=0}^{\infty} \frac{\lambda^\nu (t-s)^\nu}{\nu!} \sum_{\mu=0}^{\nu} \binom{\nu}{\mu} (-1)^{\nu-\mu} \delta_{n,n'+\mu} = \sum_{\nu=n-n'}^{\infty} \frac{\lambda^\nu (t-s)^\nu}{\nu!} \binom{\nu}{n-n'} (-1)^{\nu-(n-n')} =$$

$$\frac{(\lambda(t-s))^{n-n'}}{(n-n')!} \sum_{\nu=n-n'}^{\infty} \frac{(-\lambda(t-s))^{\nu-(n-n')}}{(\nu-(n-n'))!} = \exp[-\lambda(t-s)] \frac{(\lambda(t-s))^{n-n'}}{(n-n')!}. \quad (18)$$

Multiplying $\boldsymbol{P}(t,s)$ given in (18) by the probability distribution at the time $s = 0$, i. e. due to (10) $(n'|\boldsymbol{P}(0)) = \delta_{n',0}$, which fact simply means putting $n' = 0$ and $s = 0$, (18) is just reduced to (1), q. e. d.

It is, however, evident that the assumptions underlying our deduction of the distribution (1) are not good approximations to the real multiplication process. This fact is also clearly brought out by (3) due to which the average number should be proportional to the thickness and should thus increase indefinitely in contrast to the fact that the secondaries will become less and less energetic with increasing value of t and, therefore, finally be absorbed.

§ 4.3. The model of BHABHA and HEITLER was in fact soon criticized by FURRY[1]) who replaced it by the following. He disregards the photons and assumes that *one electron by travelling a distance Δt can be converted into two electrons with a probability asymptotically proportional to Δt, $\lambda\Delta t$.* For this stochastic process we obtain in a way analogous to (4.2.8) the equations

$$P(0,t) = 0$$

$$P(1, t+\Delta t) = P(1,t)(1-\lambda\Delta t) + o(\Delta t) \quad (1)$$

$$P(n, t+\Delta t) = P(n,t)(1-n\lambda\Delta t) + P(n-1,t)(n-1)\lambda\Delta t + o(\Delta t)$$
$$n = 2, 3, 4, \ldots$$

as now *each* electron present at t has the probability $\lambda\Delta t$ of being doubled in Δt. From (1) we obtain by making $\Delta t \to 0$ the system of simultaneous differential equations which is analogous to (4.2.9)[2])

[1]) FURRY (1937).

[2]) Comparing with the general theory in chap. 2 we see that this is a stochastic process with the intensity function $p(n,t) = n\lambda$ and the relative transition probability $(n|\boldsymbol{\Pi}|n') = \delta_{n,n'+1}$. Furthermore, we see that also the Furry process is a *pure propagation process*, the operator matrix being again a *row half matrix*. The solution is then again given by (2.6.21). The operator matrix is again constant and nearly the

$$P(0, t) \equiv 0$$

$$\frac{dP(1, t)}{dt} = -\lambda P(1, t) \tag{2}$$

$$\frac{dP(n, t)}{dt} = \lambda(n-1)P(n-1, t) - \lambda n P(n, t) \qquad n = 2, 3, 4, \ldots \quad (\lambda \geq 0, t \geq 0)$$

with the initial condition

$$P(n, 0) = \delta_{n,1}. \tag{3}$$

This system of equations governs the stochastic process which we shall call the **Furry process.** From the general theory given in chap. 2 this system (whose operator matrix is of type IV cf. § 7.6) is known to have a unique solution for which $0 \leq P(n, t) \leq 1$ and $\sum_{n=0}^{\infty} P(n, t) \equiv 1$. Also the Furry process is, however, so simple that these facts can be shown directly. On the analogy of (4.2.11) we now obtain

$$P(0, t) \equiv 0$$

$$P(1, t) = c_1 e^{-\lambda t} \tag{4}$$

$$P(n, t) = e^{-n\lambda t}\left((n-1)\lambda \int_0^t P(n-1, t')e^{n\lambda t'}dt' + c_n\right) \qquad n = 2, 3, 4, \ldots .$$

From (3) it follows that $c_1 = 1$, $c_n = 0$ and by induction it is immediately seen that (4) leads to the **Furry distribution**

$$P(0, t) \equiv 0$$

$$P(n, t) = e^{-\lambda t}(1-e^{-\lambda t})^{n-1} \qquad n = 1, 2, 3, \ldots \quad (\lambda t \geq 0). \tag{5}$$

It is easily verified that (5) really represents a probability distribution since for all $t > 0$ we have $P(n, t) > 0$ and

$$\sum_{n=0}^{\infty} P(n, t) = 0 + e^{-\lambda t} \sum_{n=1}^{\infty} (1-e^{-\lambda t})^{n-1} = e^{-\lambda t} \frac{1}{1-(1-e^{-\lambda t})} = 1. \tag{6}$$

The most characteristic feature of this distribution consists in its being monotonous in respect to n, decreasing with increasing n,

---

same as for the Poisson process, λ being only replaced by $n\lambda$. The solution of the Furry process is, consequently, again given by (2.6.7), but as the evaluation of the product-integral is now very complicated we shall refrain from giving it here.

whereas the Poisson distribution (4.2.1) is monotonously decreasing for $\lambda t < 1$ but has a maximum in the neighbourhood of $n \sim \lambda t$ for $\lambda t > 1$.

For the mean value $\bar{n}$ we find

$$\bar{n}(t) = \sum_{n=0}^{\infty} nP(n, t) = e^{\lambda t}. \tag{7}$$

Putting $1 - e^{-\lambda t} = u$ we have, namely,

$$\bar{n} = \sum_{n=0}^{\infty} n e^{-\lambda t}(1 - e^{-\lambda t})^{n-1} = e^{-\lambda t} \sum_{n=0}^{\infty} n u^{n-1} = e^{-\lambda t} \frac{d}{du} \sum_{n=0}^{\infty} u^n =$$
$$= e^{-\lambda t} \frac{d}{du} \frac{1}{1-u} = e^{-\lambda t} \frac{1}{(1-u)^2} = e^{\lambda t}. \tag{8}$$

Equation (7) can also be deduced very easily in a way analogous to (4.2.7). From the assumptions underlying the Furry process follows the *continuity equation*

$$\bar{n}(t + \Delta t) = \bar{n}(t) + \bar{n}(t)\lambda \Delta t + o(\Delta t),$$
$$\text{i. e.} \quad \frac{d\bar{n}}{dt} = \lambda \bar{n} \quad \text{whence} \quad \bar{n} = e^{\lambda t}. \tag{9}$$

For the second moment we find

$$\overline{n^2}(t) = \sum_{n=0}^{\infty} n^2 P(n, t) = 2e^{2\lambda t} - e^{\lambda t} = 2\bar{n}^2 - \bar{n}. \tag{10}$$

Putting again $1 - e^{-\lambda t} = u$ we have, namely,

$$\overline{n^2} = \sum_{n=0}^{\infty} n^2 e^{-\lambda t}(1 - e^{-\lambda t})^{n-1} = e^{-\lambda t} \sum_{n=0}^{\infty} n^2 u^{n-1} = e^{-\lambda t} \frac{d}{du} u \frac{d}{du} \sum_{n=0}^{\infty} u^n =$$
$$= e^{-\lambda t} \frac{d}{du} u \frac{d}{du} \frac{1}{1-u} = e^{-\lambda t} \frac{d}{du} \frac{u}{(1-u)^2} = e^{-\lambda t} \frac{1+u}{(1-u)^3} =$$
$$= e^{-\lambda t} \frac{2 - e^{-\lambda t}}{e^{-3\lambda t}} = 2e^{2\lambda t} - e^{\lambda t} = 2\bar{n}^2 - \bar{n}. \tag{11}$$

The absolute and relative fluctuation σ and δ for the Furry process is, therefore, given by

$$\sigma^2(t) = e^{2\lambda t} - e^{\lambda t} = \bar{n}^2 - \bar{n} \tag{12}$$

and

$$\delta = \frac{\sigma}{\bar{n}} = \left(1 - \frac{1}{\bar{n}}\right)^{\frac{1}{2}} \sim 1 \tag{13}$$

for not too small values of λt.

When applying the distribution formulae to the theory of cosmic radiation we are only interested in the distribution of the *secondary* electrons since the behaviour of the primary electron is known fairly accurately. Regarding the secondaries the Furry process can, however, be interpreted in two ways. We may either assume that the individuals do not 'die' but only give birth to *one* 'child' in each splitting process: case (a); or may assume that the individuals die giving birth to *two* 'children' in each splitting process: case (b). The total stochastic process is the same in either case but regarding the secondaries one obtains different results.

(a) In the first case we have, denoting by $P_{prim}(t)$ and $P_{sec}(n, t)$ the probability of finding the primary, respectively n secondaries, at the time t,

$$P_{prim}(t) = 1 \tag{14}$$

and thus

$$P(n, t) = P_{prim}(t) P_{sec}(n-1, t) = P_{sec}(n-1, t)\,. \tag{15}$$

From (5) we therefore obtain

$$P_{sec}(n, t) = e^{-\lambda t}(1-e^{-\lambda t})^n \qquad n = 0, 1, 2, \ldots \tag{16}$$

and, using (7),

$$\bar{n}_{sec} = \sum_{n=0}^{\infty} n P_{sec}(n, t) = (1-e^{-\lambda t})\bar{n} = \bar{n} - 1 = \bar{n} - \bar{n}_{prim} \tag{17}$$

in agreement with the well-known theorem on the mean value of a sum of two stochastic variables[1]). For $\overline{n^2_{sec}}$ we obtain, using (7) and (10),

$$\overline{n^2_{sec}} = (1-e^{-\lambda t})\overline{n^2} = 2e^{2\lambda t} - 3e^{\lambda t} + 1 = 2\bar{n}^2 - 3\bar{n} + 1 \tag{18}$$

whence

$$\sigma^2_{sec} = \overline{n^2_{sec}} - \bar{n}^2_{sec} = e^{2\lambda t} - e^{\lambda t} = \bar{n}^2 - \bar{n} = \sigma^2\,, \tag{19}$$

i. e.

$$\delta_{sec} = \frac{\sigma_{sec}}{\bar{n}_{sec}} = \left(1 - \frac{1}{\bar{n}}\right)^{-\frac{1}{2}} \sim 1 \tag{20}$$

for not too small values of λt. σ, given in (12), is here the fluctuation of the sum of primary and secondaries. (19) is besides in accordance with the well-known theorem on the fluctuation of a sum of two independent stochastic variables[2]) since it follows from (14) that $\sigma^2_{prim} = 0$.

(b) In the second case we have the primary and secondaries correlated: if at the time t we find the primary individual we can have no secondaries at all, and if we do not find the primary one we must find either 2, 3, 4, ... secondaries but we cannot find just one secondary, as illustrated in fig. 3.

1p, 0s   0p, 2s   0p, 3s
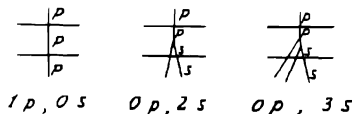
Fig. 3.

[1]) Cf. e. g. A & B p. 46.

[2]) Cf. e. g. A & B p. 47.

We thus have

$$\left.\begin{aligned}P_{sec}(0,t) &= P_{prim}(t) = P(1,t) = e^{-\lambda t}\\ P_{sec}(1,t) &\equiv 0\\ P_{sec}(n,t) &= P(n,t) = e^{-\lambda t}(1-e^{-\lambda t})^{n-1} \qquad n = 2, 3, 4, \ldots .\end{aligned}\right\} \tag{21}$$

For this distribution we have

$$\bar{n}_{prim} = e^{-\lambda t} \tag{22}$$

$$\sigma^2_{prim} = P_{prim}(1-P_{prim}) = e^{-\lambda t}(1-e^{-\lambda t}) \;^{1)} \tag{23}$$

and, using (7),

$$\bar{n}_{sec} = \sum_{n=0}^{\infty} nP_{sec}(n,t) = \sum_{n=2}^{\infty} nP(n,t) = \bar{n} - P(1,t) = \bar{n} - \bar{n}_{prim} \tag{24}$$

which formula is again in accordance with the above mentioned theorem on the mean value of a sum. Next we find, using (10),

$$\overline{n^2_{sec}} = \sum_{n=0}^{\infty} n^2 P_{sec}(n,t) = \sum_{n=2}^{\infty} n^2 P(n,t) = \overline{n^2} - P(1,t) =$$

$$2e^{2\lambda t} - e^{\lambda t} - e^{-\lambda t}\,, \tag{25}$$

and thus

$$\sigma^2_{sec} = \overline{n^2_{sec}} - \bar{n}^2_{sec} = e^{2\lambda t} - e^{\lambda t} - e^{-\lambda t} - e^{-2\lambda t} + 2\,, \tag{26}$$

i. e.,

$$\delta_{sec} = \frac{\sigma_{sec}}{\bar{n}_{sec}} = \left(\frac{1 - e^{-\lambda t} + 2e^{-2\lambda t} - e^{-3\lambda t} - e^{-4\lambda t}}{1 - 2e^{-2\lambda t} + e^{-4\lambda t}}\right)^{\frac{1}{2}} \sim 1 \tag{27}$$

for not too small values of λt.

To investigate the degree of correlation between the primary and the secondaries in this model we evaluate their correlation coefficient as defined by[2])

$$\varrho = \frac{\overline{n_{prim}\cdot n_{sec}} - \bar{n}_{prim}\cdot\bar{n}_{sec}}{\sigma_{prim}\cdot\sigma_{sec}} \qquad (-1 \leq \varrho \leq 1)\,. \tag{28}$$

We prefer, however, to evaluate it from the formula for the fluctuation of a sum of two stochastic variables[3])

$$\sigma^2 = \sigma^2_{prim} + \sigma^2_{sec} + 2\varrho\sigma_{prim}\sigma_{sec}\,. \tag{29}$$

Inserting into (29) eqs. (12), (23) and (26) we find

$$\varrho = -e^{-\frac{1}{2}\lambda t}\frac{1-e^{-2\lambda t}}{(1-2e^{-\lambda t}+3e^{-2\lambda t}-3e^{-3\lambda t}+e^{-5\lambda t})^{\frac{1}{2}}} \tag{30}$$

which shows that *the correlation is negative as might be expected in advance* since n_{prim} and n_{sec} vary in opposite directions. It is also seen that for not too small values of λt we have

[1]) Cf. e. g. A & B p. 49.

[2]) Cf. e. g. A & B p. 45.

[3]) Cf. e. g. A & B p. 47, exercise 2.

$$\varrho \sim -e^{-\frac{1}{2}\lambda t} \qquad (\lambda t \gg 1), \tag{31}$$

i. e.

$$\varrho \underset{\lambda t \to \infty}{\rightarrow} 0 \tag{32}$$

and *consequently in such case we can treat* n_{prim} and n_{sec} *as independent stochastic variables*. By expanding in Taylor's series it is finally seen that

$$\varrho \underset{\lambda t \to 0}{\rightarrow} -1 \tag{33}$$

as might be expected, for in the limit $\lambda t = 0$ we have direct dependence between n_{prim} and n_{sec}.

Comparing the relative fluctuations of the Furry process (13), (20) or (27) with the corresponding formula (4.2.6) for the Poisson process it is seen that the fluctuation about the mean value is much larger for the first than for the second distribution except for small values of λt, in which case the two distributions practically coincide. Notwithstanding the fact that the Furry process certainly furnishes a much better model than does the Poisson process it is nevertheless obvious that also this process bears only slight semblance to the real multiplication process. In both processes the important feature of the *absorption* of the electrons in the cascade, the *extinction* of the family in the biological analogy, is entirely disregarded, which fact is strongly emphasized by eqs. (4.2.3) and (7), respectively, showing that the average number of members in the electron 'family' increases linearly and exponentially, respectively, with the thickness of the layer penetrated — or the time in the biological analogy.

Next both models disregard another important feature of the real multiplication process, namely the *alternating character of the successive generations*, every second being electrons which are counted in the experiments and every second being photons which are *not* counted. Both these effects may be expected to *increase* the fluctuation because experience shows that by increasing the number of elementary processes entering into a stochastic process the fluctuation will be increased. *Our problem is, therefore, in some way or other to take account of these two features by the construction of the stochastic model.*

§ 4.4. Before we proceed with the discussion of the stochastic models of the multiplication process we shall discuss a distribution which does not in itself represent any useful model but which

belongs to the same family of distributions, embracing both the Poisson and the Furry distributions as special cases and which we shall use subsequently as an approximation in the theory of cosmic rays.

This socalled **Pólya distribution**, which is much used by statisticians in several practical problems[1]), is given by the formulae

$$P(0, t) = (1+b\lambda t)^{-\frac{1}{b}}$$

$$P(n, t) = \left(\frac{\lambda t}{1+b\lambda t}\right)^n \frac{1 \cdot (1+b) \cdots \big(1+(n-1)b\big)}{n!} P(0, t) \tag{1}$$

$$(b \geqq 0, \lambda t \geqq 0) \qquad n = 1, 2, 3, \ldots \, .$$

It is easily verified that for every fixed $\lambda t \geqq 0$ these formulae really represent a probability distribution since $P(n, t) \geqq 0$ for all $\lambda t \geqq 0$ and, because of

$$\frac{1 \cdot (1+b) \cdots (1+(n-1)b)}{n!} = (-b)^n \binom{-\frac{1}{b}}{n}, \tag{2}$$

we have

$$\sum_{n=0}^{\infty} P(n, t) =$$

$$P(0, t) \sum_{n=0}^{\infty} \binom{-\frac{1}{b}}{n} \left(\frac{-b\lambda t}{1+b\lambda t}\right)^n = (1+b\lambda t)^{-\frac{1}{b}} \left(1 - \frac{b\lambda t}{1+b\lambda t}\right)^{-\frac{1}{b}} \equiv 1\,. \tag{3}$$

It will be noticed that the Pólya distribution has *two* parameters, λt and b, in contrast to both the Poisson and the Furry distributions which have only one, λt. Next it will be seen that passing to the limit $b \to 0$ we obtain

$$P(0, t) \underset{b \to 0}{\to} e^{-\lambda t}$$

$$P(n, t) \underset{b \to 0}{\to} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \qquad n = 1, 2, 3, \ldots \tag{4}$$

which shows that the Poisson distribution (4.2.1) is in fact a special case of the Pólya distribution.

---

[1]) Cf. O. Lundberg (1940) who discusses both the theory and the different practical applications of the Pólya distribution. For the original papers cf. Pólya (1930), Eggenberger (1924), Greenwood and Yule (1920) and Newbold (1927). I wish here to express my gratitude to Dr. Lundberg for his kindness in calling my attention to the Pólya distribution.

For $b = 1$ we obtain on the other hand

$$P(0, t) = \frac{1}{1+\lambda t}$$

$$P(n, t) = \left(\frac{\lambda t}{1+\lambda t}\right)^n \frac{1}{1+\lambda t} = \frac{1}{1+\lambda t}\left(1-\frac{1}{1+\lambda t}\right)^n \qquad n = 1, 2, 3, \ldots \,. \tag{5}$$

Comparing (5) with (4.3.16) it is seen that these formulae give the same distribution as the Furry distribution of the *secondary* individuals, only with another form of the parameter. In fact, introducing into (5) a *new time scale* t' given by

$$e^{\lambda t'} = 1+\lambda t \,, \tag{6}$$

it is seen that (4.3.16) results.

For small values of b the Pólya distribution has, consequently, for each constant value of b and $\lambda t (> 1)$ a maximum with respect to n, whereas for larger values of b it is monotonous, decreasing with increasing values of n. These facts may also be seen direct as

$$\frac{P(n+1, t)}{P(n, t)} = \frac{\lambda t}{1+b\lambda t}\,\frac{1+bn}{1+n}\,, \tag{7}$$

i. e. $\qquad P(n+1, t) > P(n, t) \quad \text{for} \quad n < \lambda t(1-b)-1\,.$

In any case, if $b \geq 1$ (or if $\lambda t \leq 1$) this relation does not hold good for any value of $n (> 0)$, so the Pólya distribution is in such cases monotonously decreasing.

We shall now deduce the first and second moments of the Pólya distribution. We find, using (1), (2) and the relation

$$n\binom{x}{n} = x\binom{x-1}{n-1}, \tag{8}$$

$$\bar{n} = \sum_{n=0}^{\infty} nP(n, t) = (1+b\lambda t)^{-\frac{1}{b}} \sum_{n=0}^{\infty} n\binom{-\frac{1}{b}}{n}\left(\frac{-b\lambda t}{1+b\lambda t}\right)^n =$$

$$(1+b\lambda t)^{-\frac{1}{b}}\left(-\frac{1}{b}\right)\left(\frac{-b\lambda t}{1+b\lambda t}\right)\sum_{n=1}^{\infty}\binom{-\frac{1}{b}-1}{n-1}\left(\frac{-b\lambda t}{1+b\lambda t}\right)^{n-1} =$$

$$(1+b\lambda t)^{-\frac{1}{b}-1}(\lambda t)\left(1-\frac{b\lambda t}{1+b\lambda t}\right)^{-\frac{1}{b}-1} = \lambda t\,. \tag{9}$$

In the same way we find

$$\overline{n^2} = \sum_{n=0}^{\infty} n^2 P(n,t) = (1+b\lambda t)^{-\frac{1}{b}} \sum_{n=0}^{\infty} n^2 \binom{-\frac{1}{b}}{n} \left(\frac{-b\lambda t}{1+b\lambda t}\right)^n =$$

$$(1+b\lambda t)^{-\frac{1}{b}-1}(\lambda t) \sum_{n=1}^{\infty} [(n-1)+1] \binom{-\frac{1}{b}-1}{n-1} \left(\frac{-b\lambda t}{1+b\lambda t}\right)^{n-1} =$$

$$(1+b\lambda t)^{-\frac{1}{b}-1}(\lambda t)\left[\left(-\frac{1}{b}-1\right)\left(\frac{-b\lambda t}{1+b\lambda t}\right)\sum_{n=2}^{\infty}\binom{-\frac{1}{b}-2}{n-2}\left(\frac{-b\lambda t}{1+b\lambda t}\right)^{n-2} + \left(1-\frac{b\lambda t}{1+b\lambda t}\right)^{-\frac{1}{b}-1}\right] =$$

$$(1+b\lambda t)^{-\frac{1}{b}-1}(\lambda t)\left[(1+b)\frac{\lambda t}{1+b\lambda t}\left(1-\frac{b\lambda t}{1+b\lambda t}\right)^{-\frac{1}{b}-2} + \left(1-\frac{b\lambda t}{1+b\lambda t}\right)^{-\frac{1}{b}-1}\right] =$$

$$(1+b)(\lambda t)^2 + (\lambda t) = (1+b)\bar{n}^2 + \bar{n}\,. \tag{10}$$

For the absolute and relative fluctuations we therefore have

$$\sigma^2 = \overline{n^2} - \bar{n}^2 = b(\lambda t)^2 + \lambda t = \bar{n}(1+b\bar{n}) \tag{11}$$

and

$$\delta = \frac{\sigma}{\bar{n}} = \sqrt{b+\frac{1}{\bar{n}}} \sim \sqrt{b} \quad \left(\text{for } \bar{n} \gg \frac{1}{b}\right). \tag{12}$$

(9) and (11) show that the parameter λt is again the mean value just as in the Poisson distribution, and that the second parameter b determines direct the fluctuation which for a *constant mean value* can assume any value between $\sqrt{\bar{n}}$ and ∞. *From the fact that the two first moments can be adjusted separately comes the great applicability of the Pólya distribution in several statistical problems*[1]) *and, as discussed later, in the theory of cosmic radiation.*

It is interesting to mention that the moments (9) and (10) can also be obtained in another way, because the Pólya distribution can be interpreted as a *compound Poisson distribution*[2]). By this term is understood a mixture of several Poisson distributions with different mean values, corresponding to a certain distribution of the mean value. Let $U(x)$ be an arbitrary distribution function[3]), then a compound Poisson distribution is defined by

$$P(n,t) = \int_0^{\infty} P_{\text{Poisson}}(n, x\lambda t)\,dU(x) = \int_0^{\infty} \frac{(x\lambda t)^n}{n!} e^{-x\lambda t}\,dU(x) \tag{13}$$

$$\left(0 \leq U(x) \leq 1, \quad \int_0^{\infty} dU(x) = 1\right)$$

[1]) Cf. the literature quoted under [1]) p. 98.

[2]) Cf. LUNDBERG (1940) chap. V.

[3]) This is a function which for each value of x is equal to the probability of a stochastic variable $\boldsymbol{x}$ assuming a value $\leq x$ (cf. e. g. A & B chap. 4).

in which the integral is a STIELTJES integral[1]). Now we for the distribution of x take a Pearson curve of type III, i. e.

$$dU(x) = w_b(x)dx = \frac{\left(\frac{x}{b}\right)^{\frac{1}{b}-1}}{\left(\frac{1}{b}-1\right)!} e^{-\frac{x}{b}} d\frac{x}{b}\,, \tag{14}$$

and have from the definition of the Γ-function

$$\overline{x^k} = \int_0^\infty x^k w_b(x)dx = \frac{b^k}{\left(\frac{1}{b}-1\right)!}\int_0^\infty \left(\frac{x}{b}\right)^{k+\frac{1}{b}-1} e^{-\frac{x}{b}} d\frac{x}{b} = b^k \frac{\left(k+\frac{1}{b}-1\right)!}{\left(\frac{1}{b}-1\right)!}. \tag{15}$$

(15) gives for $k = 0$ $\quad \overline{x^0} = \int_0^\infty w_b(x)dx = 1$

$$\text{,,}\quad k = 1 \qquad \bar{x} = \int_0^\infty x w_b(x)dx = 1 \tag{16}$$

,, $k = 2$ $\quad \overline{x^2} = \int_0^\infty x^2 w_b(x)dx = 1 + b\,.$

In fact, (14) into (13) now yields

$$P(n, t) = \int_0^\infty \frac{(x\lambda t)^n}{n!} e^{-x\lambda t} \frac{\left(\frac{x}{b}\right)^{\frac{1}{b}-1}}{\left(\frac{1}{b}-1\right)!} e^{-\frac{x}{b}} d\frac{x}{b} =$$

$$\frac{(\lambda t)^n}{n!} \frac{b^n}{\left(\frac{1}{b}-1\right)!}\int_0^\infty \left(\frac{x}{b}\right)^{n+\frac{1}{b}-1} \exp\left[-\frac{1+b\lambda t}{b}x\right] d\frac{x}{b} =$$

$$\frac{(\lambda t)^n}{n!} \frac{b^n}{\left(\frac{1}{b}-1\right)!} \frac{1}{(1+b\lambda t)^{n+\frac{1}{b}}}\int_0^\infty y^{n+\frac{1}{b}-1} e^{-y} dy =$$

$$\left(\frac{\lambda t}{1+b\lambda t}\right)^n \frac{b^n}{n!} \frac{\left(n+\frac{1}{b}-1\right)!}{\left(\frac{1}{b}-1\right)!} (1+b\lambda t)^{-\frac{1}{b}} =$$

$$\left(\frac{\lambda t}{1+b\lambda t}\right)^n \frac{1\cdot(1+b)\cdots(1+(n-1)b)}{n!} (1+b\lambda t)^{-\frac{1}{b}} \tag{17}$$

which is actually the Pólya formula (1). Using (4.2.3), (4.2.4) and (16) we have at once (since everything is positive and inversion of the order of operations, therefore, is legitimate)

[1]) Cf. e. g. A & B p. 31.

$$\overline{n} = \int_0^\infty \overline{n}_{Poisson}(x) dU(x) = \lambda t \int_0^\infty x w_b(x) dx = \lambda t \tag{18}$$

in agreement with (9), and

$$\overline{n^2} = \int_0^\infty \overline{n^2}_{Poisson}(x) dU(x) = \int_0^\infty [(x\lambda t)^2 + x\lambda t] w_b(x) dx = (\lambda t)^2(1+b) + \lambda t \tag{19}$$

in agreement with (10).

The moments (9) and (10) can, finally, be obtained in a third way by a method already mentioned in the theoretical part (cf. § 2.11), which method we shall apply later. We shall, therefore, use the present simple case as an illustration of this method which consists in obtaining the moments direct from the equations governing the stochastic process *without knowing the actual solutions*. The Pólya distribution corresponds, namely, to a stochastic process of the same type as the Poisson process (§ 4.2) and the Furry process (§ 4.3), only that *the probability of one new individual being born in the 'time' interval Δt is now asymptotically put equal to*

$$\lambda \frac{1+bn}{1+b\lambda t} \Delta t\,. \tag{20}$$

Repeating the argument which led to the equations (4.2.9) and (4.3.2) we now obtain the system of simultaneous differential equations governing the **Pólya process**[1])

$$\frac{dP(0,t)}{dt} = -\frac{\lambda}{1+b\lambda t} P(0,t)$$

$$\frac{dP(n,t)}{dt} = \lambda \frac{1+b(n-1)}{1+b\lambda t} P(n-1,t) - \lambda \frac{1+bn}{1+b\lambda t} P(n,t) \qquad \begin{matrix} n = 1, 2, 3, \ldots \\ (b \geqq 0,\ \lambda \geqq 0,\ t \geqq 0) \end{matrix} \tag{21}$$

with the initial condition

$$P(n, 0) = \delta_{n,0}\,. \tag{22}$$

These equations may be solved successively in the same way as (4.3.2) was solved in (4.3.4) and in fact the Pólya distribution (1) results. In order to obtain the moments we now multiply on both sides of (21) with n^k and sum over all values of n (the convergence being ensured by the general theory in chap. 2, cf. § 2.11, as the operator matrix is of type IV, cf. § 7.6). We thus obtain

$$\begin{aligned} \frac{d}{dt}\overline{n^k} &= \sum_{n=0}^\infty n^k P'(n,t) = \\ \sum_{n=0}^\infty n^k \lambda \frac{1+b(n-1)}{1+b\lambda t} P(n-1,t) &- \sum_{n=0}^\infty n^k \lambda \frac{1+bn}{1+b\lambda t} P(n,t) = \\ &\sum_{n=0}^\infty [(n+1)^k - n^k] \lambda \frac{1+bn}{1+b\lambda t} P(n,t)\,. \end{aligned} \tag{23}$$

[1]) Comparing with the general theory in chap. 2 we see that this is a stochastic process with the intensity function $p(n,t) = \lambda \frac{1+bn}{1+b\lambda t}$ and the relative transition probability $(n|\Pi|n') = \delta_{n,n'+1}$. The Pólya process is thus also a *pure propagation process*, the solution being given by (2.6.21).

In case $k = 0$, (23) gives

$$\frac{d}{dt}\overline{n^0} = \frac{d}{dt}\sum_{n=0}^{\infty} P(n, t) = 0, \tag{24}$$

i. e., using (22),

$$\sum_{n=0}^{\infty} P(n, t) \equiv 1 \tag{25}$$

which is in agreement with (3).

In case $k = 1$, (23) gives

$$\frac{d}{dt}\bar{n} = \frac{\lambda}{1+b\lambda t}\sum_{n=0}^{\infty}(1+bn)P(n, t) = \frac{\lambda}{1+b\lambda t}(1+b\bar{n}), \tag{26}$$

or

$$\frac{d}{dt}\bar{n} - \frac{b\lambda}{1+b\lambda t}\bar{n} = \frac{\lambda}{1+b\lambda t} \qquad (\bar{n}(0) = 0) \tag{27}$$

in which the initial condition follows from (22). This is an ordinary linear differential equation of the first order and may, consequently, be integrated at once to

$$\bar{n} = (1+b\lambda t)\left(C - \frac{1}{b}\frac{1}{1+b\lambda t}\right) = \lambda t \tag{28}$$

since from the initial condition $C = \frac{1}{b}$. This agrees with (9) and (18). Equation (27) may, of course, also be obtained direct in the same way as (4.2.7) and (4.3.9). $\bar{n}(t+\Delta t)$ is, namely, a sum of the value of $\bar{n}$ at the time t and the amount created in Δt. But this last term is on our assumptions exactly $\lambda \frac{1+b\bar{n}}{1+b\lambda t}\Delta t$. We thus have the *continuity equation*

$$\bar{n}(t+\Delta t) = \bar{n}(t) + \lambda\frac{1+b\bar{n}}{1+b\lambda t}\Delta t + o(\Delta t) \tag{29}$$

which in fact leads to (27).

In case $k = 2$, (23) gives

$$\frac{d}{dt}\overline{n^2} = \sum_{n=0}^{\infty}(2n+1)\lambda\frac{1+bn}{1+b\lambda t}P(n, t) =$$

$$\frac{\lambda}{1+b\lambda t}\sum_{n=0}^{\infty}(2bn^2+(2+b)n+1)P(n, t) = \frac{\lambda}{1+b\lambda t}(2b\overline{n^2}+(2+b)\bar{n}+1), \tag{30}$$

or, using (28),

$$\frac{d}{dt}\overline{n^2} - \frac{2b\lambda}{1+b\lambda t}\overline{n^2} = \frac{\lambda}{1+b\lambda t}((2+b)\lambda t+1) \qquad (\overline{n^2}(0) = 0) \tag{31}$$

in which the initial condition follows from (22). This is an equation of the same simple type as (27) and its solution is

$$\overline{n^2} = (1+b\lambda t)^2\left(C - \frac{1}{b^2}\frac{2+b}{(1+b\lambda t)} + \frac{1}{b^2}\frac{1}{(1+b\lambda t)^2}\right) = (1+b)(\lambda t)^2 + \lambda t \tag{32}$$

since from the initial condition $C = \frac{1+b}{b^2}$. This agrees with (10) and (19).

Before we leave the Pólya distribution we may note that if the mean value $\lambda t \gg 1$ [1]) we have an asymptotic formula corresponding to the well-known Laplace formula for the Poisson distribution. We shall not here give an exact deduction, which e. g. could be obtained by using the Stirling approximation formula on the factorials in (1), but shall only remark that if $\lambda t \gg 1$, the function

$$f(x) = \lambda t \frac{(x\lambda t)^n e^{-x\lambda t}}{n!} \qquad \left(\int_0^\infty f(x)dx = 1\right) \tag{33}$$

will have a very sharp maximum $f_{max} = \lambda t \frac{n^n e^{-n}}{n!}$ for $x_{max} = \frac{n}{\lambda t}$. The main contribution to the integral in (13) will thus come from the neighbourhood of this point and as a first approximation we may, therefore, put

$$P(n, t) \sim w_b(x_{max}) \frac{1}{\lambda t}\int_0^\infty \frac{(x\lambda t)^n e^{-x\lambda t}}{n!} d(x\lambda t) = \frac{1}{b\lambda t}\frac{1}{\left(\frac{1}{b}-1\right)!}\left(\frac{n}{b\lambda t}\right)^{\frac{1}{b}-1} e^{-\frac{n}{b\lambda t}}, \quad (\lambda t \gg 1) \tag{34}$$

a result which is in fact also obtained by a more exact treatment. This formula is very handy whereas the exact formula (1) is extremely cumbersome for large values of λt, i. e. of n. In fact it follows from (34) that for $\lambda t \gg 1$ we can put

$$P(\mathbf{n} \geq n, t) = \sum_n^\infty{}' P(n, t) \sim \int_n^\infty \frac{1}{\left(\frac{1}{b}-1\right)!}\left(\frac{x}{b\lambda t}\right)^{\frac{1}{b}-1} \exp\left[-\frac{x}{b\lambda t}\right] d\left(\frac{x}{b\lambda t}\right) \tag{35}$$

which integral is tabulated as a function of $\frac{n}{b\lambda t}$ and $\frac{1}{b}$ [2]).

§ 4.5. Before giving our generalization of the Furry process, announced at the end of § 4.3, we shall very shortly report three papers dealing with our subject[3]).

The first paper by Euler does not give any stochastic model but only discusses the fluctuation in a family of generations when the fluctuation in each generation is given by the Poisson formula $\sigma = \sqrt{n}$. He therefore does not make any contribution to the problem of the *form* of the distribution but simply assumes this to

[1]) This case will not play any important rôle in our applications. In the theory of very large showers, the socalled *Hoffmannstösse*, it would, however, predominate and then our asymptotic formula would be indispensable.

[2]) Cf. Pearson (1922).

[3]) Euler (1938). Cf. especially p. 465–470.
Nordsieck, Lamb and Uhlenbeck (1940).
Arley (1938). Cf. especially p. 526–529.

be Gauss'ian (as he is interested only in cases with large mean values). He finds, in agreement with our results, that the fluctuation is, in fact, much larger than the Poisson fluctuation.

The next paper is by UHLENBECK and his collaborators. They give a generalization of the Furry process taking into account that the energy of the primary electron is divided between all its descendant secondaries. They thus obtain a very complicated stochastic process with *two* continuous parameters, the thickness of the layer and the energy of the members of the electron 'family'. Only under very simplifying assumptions regarding the probabilities of the elementary processes do they succeed in calculating the fluctuation in terms of the mean value, but are unable to make any assertions regarding the probabilities themselves. Their result is, in accordance with Euler's results and the results we shall obtain later, to the effect that the fluctuation is much larger than the Poisson fluctuation, $\sigma = \sqrt{\bar{n}}$, and that even in the tail of the multiplication curve, $\bar{n}(t)$, there are large fluctuations, the approach to the Poisson formula being very slow.

The third paper is by the author himself. We there discussed the following stochastic model: we assumed, firstly, that an electron by travelling a distance Δt has the probability Δt of being absorbed by emitting two light quanta. Secondly, we assumed that a light quantum by travelling a distance Δt has the same probability Δt of being absorbed creating two electrons. In order to account for the energy depreciation taking place in each such step, we thirdly assumed that when *three* such transformations have occurred, the resulting light quanta have lost so much energy that they can no longer create electrons and that, consequently, no further multiplication takes place. Denoting by $P(n, m, t)$ (10 such functions in all) the probability that $n(\leq 4)$ electrons and $m(\leq 8)$ light quanta are present at t, it is easy to write down and solve the differential equations governing this *two-dimensional* stochastic process. It is then found that the Poisson distribution is a very close approximation to the *marginal distribution* $P(n, t) = \sum_m P(n, m, t)$ giving the probability of n electrons being present, independent of the number of light quanta found. This fact is, however, not to be wondered at. The mean value of this process is, namely, found to be

$$\bar{n} = \sum_{n=0}^{4} nP(n, t) = (1+2t^2)e^{-t} \tag{1}$$

and the maximum value of this function is 1·24 (for $t = 1+\frac{1}{\sqrt{2}} =$ 1·707). But if the mean value of a stochastic process is of the order one, the deviations from a Poisson distribution, or any other distribution, with the same mean value will, of course, be negligible. *The author's problem has, therefore, been to generalize the process to give arbitrarily high values of the mean value.* An obvious idea would be to alter the third assumption so as to allow of more transformations taking place before we regard the particles as being 'dead' or at any rate 'sterile'. This is, however, not a good idea. It will, namely, be seen that the process discussed is not what is called in the theoretical part, chap. 1, a *stochastically definite* since the behaviour of the individuals depends not solely on the number present but also on their individual antecedents. This fact makes it impossible to write down *general* equations for $P(n, m, t)$ and the generalized process is, therefore, *in practice* useless for higher mean values.

§ **4.6.** We now return to the discussion at the end of § 4.3. As we saw there the problem was to generalize the Furry process in such a way that we incorporate both the fact that the individuals in the process are absorbed, i. e. 'die' in the biological analogy, and that we are only interested in every second generation. *We begin with the absorption.* The first idea to start with in order to take account of the absorption would be simply to *assume each individual to have the asymptotic probability $\gamma\Delta t$ of 'dying' during the time interval Δt.* Instead of the equations (4.3.1) we should thus have

$$P(0, t+\Delta t) = P(0, t)+P(1, t)\gamma\Delta t+o(\Delta t)$$

$$P(n, t+\Delta t) = P(n+1, t)(n+1)\gamma\Delta t+ \qquad (1)$$

$$P(n, t)\big(1-n(\lambda+\gamma)\Delta t\big)+P(n-1, t)(n-1)\lambda\Delta t+o(\Delta t)$$

$$n = 1, 2, 3, \ldots$$

since the result n individuals at the time $t+\Delta t$ may be obtained in three different ways: either $n+1$ present at the time t and one dying during Δt, or n present at t and neither 'birth' nor 'death' taking place during Δt or, finally, $n-1$ present at t and one being born during Δt. Dividing by Δt and making the limit $\Delta t \to 0$ we obtain from (1) the system of simultaneous differential equations[1])

[1]) Comparing with the general theory in chap. 2 we see that this is a stochastic process with the intensity function $p(n, t) = (\lambda+\gamma)n$ and the relative transition

$$\frac{dP(0, t)}{dt} = \gamma P(1, t)$$
$$\frac{dP(n, t)}{dt} = \gamma(n+1)P(n+1, t)-(\lambda+\gamma)nP(n, t)+\lambda(n-1)P(n-1, t) \tag{2}$$

with the initial condition $n = 1, 2, 3, \ldots \ (\gamma \geqq 0, \lambda \geqq 0, t \geqq 0)$

$$P(n, 0) = \delta_{n, 1}. \tag{3}$$

For $\gamma \to 0$ this system will just reduce to (4.3.2), but for $\gamma > 0$ it is already so complicated that the equations cannot be solved elementarily since they do not allow of a successive solution like (4.3.2), each $P(n, t)$ depending now not only on the *lower* but also on the *higher* $P(n, t)$-functions.

For the evaluation of the probabilities we must thus have recourse to the *product-integral* (cf. chap. 2). As the operator matrix is constant the solution is simply given by (2.6.7), i. e.

$$\boldsymbol{P}(t, s) = \exp\left[\boldsymbol{A}(t-s)\right] \tag{4}$$

with $\boldsymbol{A}$, due to (2), given by

$$\boldsymbol{A} = \left\{\begin{matrix} 0 & \gamma \cdot 1 & 0 & 0 & \ldots \\ 0 & -(\lambda+\gamma)\cdot 1 & \gamma \cdot 2 & 0 & \ldots \\ 0 & \lambda \cdot 1 & -(\lambda+\gamma)\cdot 2 & \gamma \cdot 3 & \ldots \\ 0 & 0 & \lambda \cdot 2 & -(\lambda+\gamma)\cdot 3 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix}\right\}. \tag{5}$$

The powers of A may be worked out as far as necessary although this calculation, of course, will grow more and more cumbersome the higher the power. It will be seen that in this way we obtain the solutions in the form of *power series* in $t-s$. From theorem 1 § 7.4 it follows that these series are absolutely and uniformly convergent for such values of $t-s$ as are limited by the condition $t-s < \frac{1}{2(\lambda+\gamma)}$ (cf. (7.1.8) and (7.4.2)-(7.4.4)). As shown

---

probability $(n'+1|\boldsymbol{\Pi}|n') = \frac{\lambda}{\lambda+\gamma}$, $(n'-1|\boldsymbol{\Pi}|n') = \frac{\gamma}{\lambda+\gamma}$ and all other values equal to 0. Furthermore, we see that the process is a *generalized elementary process* (cf. § 2.9) characterized by the operator matrix being *column semi-diagonal* with $l = 1$ (cf. § 7.3). Next, the condition (2.9.19) is satisfied, i. e. the intensity function divided by n is uniformly bounded and the operator matrix is thus of type IV (cf. § 7.6). As shown in § 2.9 the process then satisfies the FELLER-LUNDBERG condition for the total 'probability mass' remaining in finity for all times, i. e. $\sum_{n=0}^{\infty} P(n, t) = 1$. Finally, all the moments exist as shown in § 2.11 and satisfy (2.11.16).

in § 2.7 the solution may, however, be obtained for arbitrarily high values of t by *exponentiable continuation*.

The stochastic process (2) has been discussed by FELLER[1]) in an interesting paper on the application of stochastic processes in biology. It is, however, easily seen that it is no better model in our case. For the mean value we have, namely, as a generalization of (4.3.9) the *continuity equation*

$$\bar{n}(t+\Delta t) = \bar{n}(t)+\bar{n}(t)\lambda\Delta t-\bar{n}\gamma\Delta t+o(\Delta t),$$

i. e.

$$\frac{d\bar{n}}{dt} = (\lambda-\gamma)\bar{n} \tag{6}$$

whence

$$\bar{n} = e^{(\lambda-\gamma)t} \tag{7}$$

since from (3) $\bar{n}(0) = 1$. (7) shows that for $\lambda > \gamma$ the mean value will increase monotonously, for $\lambda = \gamma$ be constant and for $\lambda < \gamma$ decrease monotonously, whereas we want the mean value to pass through a *maximum* (cf. chap. 5).

Before passing to the discussion of a better model, we shall, however, point out an interesting feature of this stochastic process. *Including absorption means, namely, increasing the fluctuation* which is in accordance with the fact previously noted that the more elementary processes in which the individuals can participate, the greater the fluctuation.

The general moments can easily be obtained from (2) by the method illustrated in the case of the Pólya distribution (cf. the end of § 4.4). Multiplying with n^k on both sides and summing over all values of n we obtain

$$\frac{d}{dt}\overline{n^k} = \sum_{n=0}^{\infty} n^k P'(n, t) =$$

$$\sum_{n=0}^{\infty} n^k\gamma(n+1)P(n+1, t) - \sum_{n=0}^{\infty} n^k(\lambda+\gamma)nP(n, t) + \sum_{n=0}^{\infty} n^k\lambda(n-1)P(n-1, t) =$$

$$\sum_{n=0}^{\infty}\Big[((n+1)^k-n^k)\lambda-(n^k-(n-1)^k)\gamma\Big]nP(n, t). \tag{8}$$

This equation shows, since it contains only the moments of *lower* order, that all the moments can be calculated *successively*.

In case $k = 0$, (8) gives

$$\frac{d}{dt}\overline{n^0} = \frac{d}{dt}\sum_{n=0}^{\infty} P(n, t) = 0, \tag{9}$$

i. e., using (3),

$$\sum_{n=0}^{\infty} P(n, t) \equiv 1. \tag{10}$$

[1]) FELLER (1939).

In case $k = 1$, (8) gives

$$\frac{d}{dt}\bar{n} = \sum_{n=0}^{\infty}(\lambda-\gamma)nP(n, t) = (\lambda-\gamma)\bar{n} \tag{11}$$

which is just (6).

In case $k = 2$, (8) gives

$$\frac{d}{dt}\overline{n^2} = \sum_{n=0}^{\infty}[(2n+1)\lambda-(2n-1)\gamma]nP(n, t) = 2(\lambda-\gamma)\overline{n^2}+(\lambda+\gamma)\bar{n} \tag{12}$$

or, using (7),

$$\frac{d}{dt}\overline{n^2}-2(\lambda-\gamma)\overline{n^2} = (\lambda+\gamma)e^{(\lambda-\gamma)t} \qquad (\overline{n^2}(0) = 1) \tag{13}$$

in which the initial condition follows from (3). The solution of (13) is, using the initial condition,

$$\overline{n^2} = \begin{cases} e^{2(\lambda-\gamma)t}\left(\dfrac{2\lambda}{\lambda-\gamma}-\dfrac{\lambda+\gamma}{\lambda-\gamma}e^{-(\lambda-\gamma)t}\right) & (\lambda \neq \gamma) \\ 1+2\lambda t & (\lambda = \gamma)\,. \end{cases} \tag{14}$$

So

$$\sigma^2 = \begin{cases} \dfrac{\lambda+\gamma}{\lambda-\gamma}e^{2(\lambda-\gamma)t}(1-e^{-(\lambda-\gamma)t}) = \dfrac{\lambda+\gamma}{\lambda-\gamma}(\bar{n}^2-\bar{n}) & (\lambda \neq \gamma) \\ 2\lambda t & (\lambda = \gamma) \end{cases} \tag{15}$$

which shows that the fluctuation is in fact greater with than without absorption. Especially, we note the interesting fact that if $\lambda = \gamma$, which means that the mean value is a constant equal to 1, *the fluctuation about this constant mean value will increase indefinitely with the square root of the time!*

§ 4.7. In our attempt to replace the stochastic process just discussed by a better one we are guided by the knowledge of the *form* desired for the mean value. As we shall recapitulate in the next chapter the mean value of electrons (and also of photons) may be calculated from the quantum theory and the result will be curves of the form shown in fig. 4 which give the average number of electrons in a cascade shower produced by one 'parent' photon as a function of the thickness of the material traversed (cf. table 13 p. 152). For a 'parent' electron the curves would be similar (cf. table 15 p. 159). It will be seen that such curves may be fairly well approximated by Gauss'ian error curves (although their tail falls off more slowly than does the tail of a Gauss'ian curve). This feature is, however, very easily taken into account simply by *assuming that the probability of an individual dying during the time interval Δt is not*
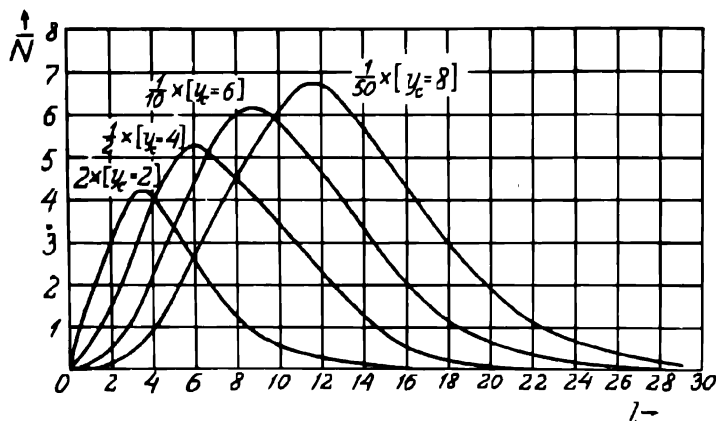
FIG. 4. *The mean number of electrons in a cascade shower, $\bar{N}$ (cf. table 13), as a function of the thickness of the material traversed, in shower-units l (given in (5.2.5)), for photon-initiated showers corresponding to various values of the primary energy parameter, y_c (given in (5.2.21)).*

—*asymptotically*—$\gamma\Delta t$ *as previously but is proportional to the 'age' of the whole family, i.e.* $\gamma t\Delta t$. This assumption is also quite plausible physically since *on the average* the particles in the cascade will have less and less energy with increasing values of t, the thickness of the material traversed and, thus, *on the average* have greater and greater probability of being absorbed. *But we stress that our assumption is, of course, an idealization of the real multiplication process, since in this the probability of being absorbed—and moreover also of giving birth to new particles—depends on the energy of the particle at that time and this energy, which depends again on the antecedents of each individual electron or photon, will fluctuate around the average energy. We think, however, that our assumption is the one bearing the closest resemblance to actual events and being at the same time sufficiently simple* (*cf.* p. 142).

In fact, it is easily seen that our new assumption really gives the desired form of the mean value curves. Instead of (4.6.6) we now have the *continuity equation*

$$\bar{n}(t+\Delta t) = \bar{n}(t)+\bar{n}(t)\lambda\Delta t-\bar{n}\gamma t\Delta t+o(\Delta t),$$

i.e.

$$\frac{d\bar{n}}{dt} = (\lambda-\gamma t)\bar{n} \tag{1}$$

whence

$$\bar{n} = \exp\left[\lambda t - \frac{\gamma}{2}t^2\right] = \exp\left[-\frac{\gamma}{2}\left(t - \frac{\lambda}{\gamma}\right)^2 + \frac{\lambda^2}{2\gamma}\right] \tag{2}$$

since from the initial condition $\bar{n}(0) = 1$. (2) shows that the mean value curves are Gauss'ian curves with the maximum value

$$\bar{n}_{max} = \exp\left[\frac{\lambda^2}{2\gamma}\right] \quad \text{for} \quad t_{max} = \frac{\lambda}{\gamma}. \tag{3}$$

We note that we could just as easily obtain mean value curves with tails falling off more slowly than the tail of the Gauss'ian curves if we assumed the 'death-rate' probability to be of the form $\gamma t^{\varepsilon} \Delta t$ where $0 \leq \varepsilon \leq 1$, e.g. $\varepsilon = \frac{1}{2}$. Instead of (1) and (2) we should then have obtained

$$\frac{d\bar{n}}{dt} = (\lambda - \gamma t^{\varepsilon}) \tag{4}$$

whence

$$\bar{n} = \exp\left[\lambda t - \frac{\gamma}{\varepsilon + 1} t^{\varepsilon+1}\right]. \tag{5}$$

This formula can give any asymmetry desired by the proper adjustment of the parameter ε. As we shall see later this assumption would be mathematically much less handy (by the evaluation of the fluctuation in the application we are going to make of the model, cf. chap. 6) and we think the improvement obtained would not compensate for this complication. We shall, therefore, keep to the assumption $\varepsilon = 1$.

Instead of (4.6.2) we now find[1])

$$\frac{dP(0, t)}{dt} = \gamma t P(1, t) \tag{6}$$

$$\frac{dP(n, t)}{dt} = \gamma t(n+1)P(n+1, t) - (\lambda + \gamma t)nP(n, t) + \lambda(n-1)P(n-1, t)$$
$$n = 1, 2, 3, \ldots \quad (\gamma \geq 0, \lambda \geq 0, t \geq 0)$$

with the same initial condition

$$P(n, 0) = \delta_{n,1}. \tag{7}$$

[1]) Comparing with the general theory in chap. 2 we see that this is a stochastic process with the intensity function $\mu(n, t) = (\lambda + \gamma t)n$ and the relative transition probability $(n'+1|\Pi|n') = \frac{\lambda}{\lambda + \gamma t}$, $(n'-1|\Pi|n') = \frac{\gamma t}{\lambda + \gamma t}$ and all other values equal to 0. Just as the model in § 4.6 the present model is a *generalized elementary process* with $l = 1$ (cf. § 2.9), the operator matrix being of type IV (cf. § 7.6). All the moments thus exist (cf. § 2.11) and satisfy (2.11.16), and especially we have $\overline{n^0} = \sum_{n=0}^{\infty} P(n, t) = 1$.

The author has spent much time in trying to obtain the solutions of this infinite system of simultaneous differential equations expressed in a closed form but has not succeeded.

We may, naturally, have recourse to the product-integral which always offers an exact solution, but by a closed form we mean a form in which the series of iteration (2.4.1) has been summed. In the present case the operator matrix $\boldsymbol{A}(t)$, given in (4.6.5) with γ replaced by γt, does *not* commute with its integral, $\boldsymbol{F}_1 = \int_s^t \boldsymbol{A} dt$, as e. g. $(0|\boldsymbol{A} \cdot \boldsymbol{F}_1|1) = -\gamma t \left(\lambda(t-s) + \frac{\gamma}{2}(t^2-s^2)\right) \neq$ $(0|\boldsymbol{F}_1 \cdot \boldsymbol{A}|1) = -(\lambda+\gamma t)\frac{\gamma}{2}(t^2-s^2)$. Consequently, the product-integral cannot be expressed in the simple form (2.6.7) as was the case with the model in § 4.6. The series (2.4.1) may, however, also be easily worked out as far as necessary and the result will again be a solution expressed as a power series, but this time in t and not in $t-s$ as for a *constant* operator matrix. Furthermore, in this case the resulting series will only be convergent for such values of t as are limited by the condition $t-s < \frac{1}{2(\lambda+\gamma t)}$ (cf. (7.1.8) and (7.4.2)–(7.4.4)). As shown in example 1 § 2.7 the solution may, however, still be obtained for arbitrarily high values of t by *exponentiable continuation* (cf. § 2.7).

Even in the simple case of (4.6.2), in which the coefficients were constant, it seems impossible to reduce the solution to simple quadrature as was the case with the equations we have so far encountered (cf. (4.2.9), (4.3.2) and (4.4.21)). The reason is, as already mentioned, that the n'th function depends not only on the *lower* functions, as is the case in stochastic processes *without* absorption, but also on the *higher* ones, which fact precludes the possibility of a successive solution. Even if, however, it should be possible to find closed expressions for the solution, which at any rate from the general theory in chap. 2 is known to exist, to be unique and to form a probability distribution, these expressions would probably be very complicated functions both of t, λ and γ. Since the parameters λ and γ determine the maximum of the mean value they have to be correlated with the energy of the primary individual in the real multiplication process. Applying later the probability distribution obtained here to the theory of cosmic rays we would have, as shown in chap. 6, to integrate over the parameters λ and γ which calculation would certainly be far too complicated. We should, consequently, even in the case of exact solutions expect to be forced to have recourse to approximations to the probability distribution.

We shall, however, refrain from discussing such approximations at this stage since we still have to generalize our process in such a way that account is taken of the fact that only every second generation is of interest Before we give this generalization we shall, however, deduce the equation for the moments. Instead of (4.6.8),

and since we only have to make the transformation $\gamma \to \gamma t$, we now obtain

$$\frac{d}{dt}\overline{n^k} = \sum_{n=0}^{\infty} n^k P'(n, t) = \sum_{n=0}^{\infty} \left[\left((n+1)^k - n^k\right)\lambda - \left(n^k - (n-1)^k\right)\gamma t\right] n P(n, t) . \quad (8)$$

In case $k = 0$, (8) gives as usual

$$\frac{d}{dt}\overline{n^0} = \frac{d}{dt}\sum_{n=0}^{\infty} P(n, t) = 0 , \quad (9)$$

i. e., using (7),

$$\sum_{n=0}^{\infty} P(n, t) \equiv 1 . \quad (10)$$

In case $k = 1$, (8) gives

$$\frac{d}{dt}\bar{n} = \sum_{n=0}^{\infty} (\lambda - \gamma t) n P(n, t) = (\lambda - \gamma t)\bar{n} \quad (11)$$

which is just (1).

In case $k = 2$, (8) gives

$$\frac{d}{dt}\overline{n^2} = \sum_{n=0}^{\infty}[(2n+1)\lambda - (2n-1)\gamma t] n P(n, t) = 2(\lambda - \gamma t)\overline{n^2} + (\lambda + \gamma t)\bar{n} \quad (12)$$

or, using (2),

$$\frac{d}{dt}\overline{n^2} - 2(\lambda - \gamma t)\overline{n^2} = (\lambda + \gamma t) \exp\left[\lambda t - \frac{\gamma}{2} t^2\right] \quad (\overline{n^2}(0) = 1) \quad (13)$$

in which the initial condition follows from (7). Using the initial condition, the solution of (13) is

$$\overline{n^2} = \exp\left[\lambda t - \frac{\gamma}{2} t^2\right] + 2\lambda \exp\left[2\left(\lambda t - \frac{\gamma}{2} t^2\right)\right] \int_0^t \exp\left[-\lambda t' + \frac{\gamma}{2} t'^2\right] dt' =$$

$$= \bar{n} + 2\lambda \bar{n}^2 \int_0^t \exp\left[-\lambda t' + \frac{\gamma}{2} t'^2\right] dt' . \quad (14)$$

For $\gamma \to 0$ this formula just reduces to (4.3.10).

For numerical evaluation of (14) the function tabulated in chap. 8

$$\psi(x) = -\psi(-x) = \int_0^x \exp[t^2] dt \quad (15)$$

may be used, and then we may write

$$\overline{n^2} = \bar{n} + 2\lambda \bar{n}^2 \sqrt{\frac{2}{\gamma}} \exp\left[-\frac{\lambda^2}{2\gamma}\right]\left(\psi\left(\frac{\lambda}{\sqrt{2\gamma}}\right) + \psi\left(\sqrt{\frac{\gamma}{2}}\, t - \frac{\lambda}{\sqrt{2\gamma}}\right)\right) . \quad (16)$$

We see here that if we had not chosen $\varepsilon = 1$ (cf. p. 111, especially (4.7.5)) we had been unable to reduce the integral in (14) to one and the same function, $\psi(x)$, independently of γ, but would have had to calculate the integral separately for each value of γ. This fact would render the practical use of the stochastic process very cumbersome (cf. chap. 6).

§ 4.8. We shall now accomplish the final generalization of the Furry process. *Having taken the absorption into account in the previous paragraph we now split the generations in the model discussed there into the generations of even and odd numbers.* If the primary particle is an electron, the generations of *even* number will also be electrons and those of *odd* number will be photons, and vice versa if the primary particle is a photon. We thus have to look for the probability of finding n individuals of the one kind—which we shall simply call n-particles—and m of the other kind—m-particles—at the 'time' t. We thus have a *two-dimensional* stochastic process. The theory of such processes we have discussed in chap. 3. In order to simplify the calculations we assume the process to be *symmetric* as regards the two kinds of particles, apart, of course, from the asymmetry introduced by the initial condition. More exactly, we make the following assumptions: *particles of each kind have a certain probability of being absorbed in a time interval Δt at the same time giving birth to two particles of the other kind. This probability is asymptotically given by $\lambda\Delta t$ for both n- and m-particles.* Moreover, *each particle has a probability asymptotically given by $\gamma t\Delta t$—for both n- and m-particles—of being absorbed during Δt without giving birth to any new particles.* The parameter λ is thus a measure of the 'birth-rate', γ of the 'death-rate' of the 'family'. With these assumptions we have as a generalization of (4.7.1) the *continuity equations* for the mean values

$$\bar{n}(t+\Delta t) = \bar{n}(t) + \bar{m}(t)2\lambda\Delta t - \bar{n}(t)\lambda\Delta t - \bar{n}(t)\gamma t\Delta t + o(\Delta t)$$

$$\bar{m}(t+\Delta t) = \bar{m}(t) + \bar{n}(t)2\lambda\Delta t - \bar{m}(t)\lambda\Delta t - \bar{m}(t)\gamma t\Delta t + o(\Delta t)\,,$$

i. e.

$$\frac{d\bar{n}}{dt} = -(\lambda+\gamma t)\bar{n} + 2\lambda\bar{m}$$
$$\frac{d\bar{m}}{dt} = -(\lambda+\gamma t)\bar{m} + 2\lambda\bar{n}\,. \tag{1}$$

As an initial condition we assume one n-particle and no m-particles present at the time $t = 0$, i. e.

$$\bar{n}(0) = 1 \quad \text{and} \quad \bar{m}(0) = 0\,. \tag{2}$$

For the total number of both n- and m-particles (1) gives by simple addition, using a well-known theorem on mean values[1])

$$\frac{d}{dt}(\overline{n+m}) = (\lambda - \gamma t)(\overline{n+m}) \tag{3}$$

in agreement with (4.7.1). In order to solve (1) we express m by the first equation and introduce the result in the second one. We thus find

$$\bar{m} = \frac{1}{2\lambda}\left(\frac{d}{dt}\bar{n} + (\lambda + \gamma t)\bar{n}\right) \tag{4}$$

and

$$\frac{d^2\bar{n}}{dt^2} + 2(\lambda + \gamma t)\frac{d\bar{n}}{dt} + [(\lambda + \gamma t)^2 + \gamma - 4\lambda^2]\bar{n} = 0\,. \tag{5}$$

This differential equation of the second order cannot be solved by any standard method. It is, however, a natural idea that it must have a particular integral of the form (4.7.2). In fact it is easily verified that

$$\varphi_1(t) = \exp\left[\lambda t - \frac{\gamma}{2}t^2\right] \tag{6}$$

is a solution. From the ordinary theory of differential equations of the second order another solution is then found to be

$$\varphi_2(t) = \exp\left[-3\lambda t - \frac{\gamma}{2}t^2\right]. \tag{7}$$

The total integral of (5) is, consequently, $c_1\varphi_1 + c_2\varphi_2$ and, using (2) and (4), we finally have

$$\begin{aligned} \bar{n} &= \frac{1}{2}\exp\left[\lambda t - \frac{\gamma}{2}t^2\right] + \frac{1}{2}\exp\left[-3\lambda t - \frac{\gamma}{2}t^2\right] \\ \bar{m} &= \frac{1}{2}\exp\left[\lambda t - \frac{\gamma}{2}t^2\right] - \frac{1}{2}\exp\left[-3\lambda t - \frac{\gamma}{2}t^2\right]. \end{aligned} \tag{8}$$

These formulae show that for not too small values of λt we have $\bar{n} \sim \bar{m}$ and that, therefore, each one is equal to half the old mean value in the one-dimensional model. This fact might be expected since the asymmetry introduced by the initial condition must be less and less pronounced the larger λt becomes. We seek the maximum of $\bar{n}$ and $\bar{m}$ and find by differentiation the condition

$$\frac{\lambda - \gamma t}{3\lambda + \gamma t} = \pm \exp[-4\lambda t] \tag{9}$$

[1]) Cf. e. g. A & B p. 46.

for $\bar{n}$ and $\bar{m}$, respectively. If, for the corresponding value of t, t_{max}, we have $\lambda t_{max} > 1$, the exponential is approximately equal to 0 and (9) gives the same value as (4.7.3)

$$\bar{n}_{max} = \bar{m}_{max} = \frac{1}{2} \exp\left[\frac{\lambda^2}{2\gamma}\right] \quad \text{for} \quad t_{max} = \frac{\lambda}{\gamma} \left(\lambda t_{max} = \frac{\lambda^2}{\gamma} > 1\right). \quad (10)$$

The other term is, namely, equal to $\exp\left[-3\frac{\lambda^2}{\gamma} - \frac{\gamma}{2}\frac{\lambda^2}{\gamma^2}\right] = \exp\left[-\frac{7}{2}\frac{\lambda^2}{\gamma}\right] < \exp\left[-\frac{7}{2}\right] = 0{\cdot}03$ which is quite negligible compared with the former term.

Regarding the fluctuations it may very easily be seen without explicitly evaluating them that the splitting of the one-dimensional model into a two-dimensional model will result in *increasing* the fluctuation. As we have just seen, the asymmetry from the initial condition becomes negligible for not too small values of λt. We therefore have $\bar{n} \sim \bar{m} \sim \frac{\overline{n+m}}{2}$ and $\sigma_n \sim \sigma_m$ for such values of t. Since[1])

$$\sigma^2_{n+m} = \sigma^2_n + \sigma^2_m + 2\varrho\sigma_n\sigma_m \sim 2\sigma^2_n(1+\varrho) \,, \quad (11)$$

we have for the relative fluctuation

$$\frac{\sigma_n}{\bar{n}} \sim \frac{\sigma_{n+m}}{\sqrt{2(1+\varrho)}} \frac{1}{\frac{\overline{n+m}}{2}} = \sqrt{\frac{2}{1+\varrho}} \frac{\sigma_{n+m}}{\overline{n+m}} \geq \frac{\sigma_{n+m}}{\overline{n+m}} \,, \quad (12)$$

because for the correlation coefficient we have $|\varrho| \leq 1$. (12) shows that the fluctuation is in fact increased by splitting the one-dimensional model into a two-dimensional model.

Before deducing the exact expressions for the quadratic moments we have to set down the differential equations governing our stochastic process. From our assumptions we obtain, denoting by $P(n, m, t)$ the probability that at the 'time' t there are present n particles of the first kind and m of the second kind,

$$P(n, m, t+\Delta t) = P(n+1, m, t)(n+1)\gamma t\Delta t + P(n, m+1, t)(m+1)\gamma t\Delta t +$$
$$P(n, m, t)\big(1-(n+m)(\lambda+\gamma t)\Delta t\big) +$$
$$P(n+1, m-2, t)(n+1)\lambda\Delta t + P(n-2, m+1, t)(m+1)\lambda\Delta t + o(\Delta t) \quad (13)$$

$(P(n, m, t) = 0$ if one or both indices become negative$)$.

[1]) Cf. e. g. A & B p. 47, exercise 2.

The first two terms correspond to the 'death' of one of either particles during Δt, the last two terms to the 'birth' of two new particles (the parent-particle thereby 'dying') and the middle term corresponds to the third eventuality of neither 'death' nor 'birth' taking place during Δt. From (13) we obtain as a generalization of (4.7.6) by making $\Delta t \to 0$ the double-infinite system of simultaneous differential equations[1])

$$\frac{dP(n, m, t)}{dt} = (n+1)\gamma t P(n+1, m, t) + (m+1)\gamma t P(n, m+1, t) -$$
$$(n+m)(\lambda+\gamma t)P(n, m, t) + (n+1)\lambda P(n+1, m-2, t) +$$
$$(m+1)\lambda P(n-2, m+1, t) \qquad n, m = 0, 1, 2, \dots \tag{14}$$

$(P(n, m, t) \equiv 0 \quad \text{for} \quad n \text{ or } m = -1, -2, -3, \dots) \quad (\gamma \geqq 0, \lambda \geqq 0, t \geqq 0)$

with the initial condition

$$P(n, m, 0) = \delta_{n,1} \cdot \delta_{m,0}. \tag{15}$$

It is easily verified that $\sum_{n+m=r}' P(n, m, t)$ satisfies the old equation (4.7.6) as should be the case, since we have now only split the old process into even and odd generations. In fact, putting $Q(r, t) = \sum_{n+m=r}' P(n, m, t) = \sum_{n=0}^{r}' P(n, r-n, t)$, we have

$$\sum_{n+m=r}' [(n+1)P(n+1, m) + (m+1)P(n, m+1)] =$$
$$\sum_{n=0}^{r}' [(n+1)P(n+1, r-n) + (r-n+1)P(n, r-n+1)] =$$
$$\sum_{n'=0}^{r+1}' n'P(n', r-n'+1) + \sum_{n=0}^{r+1}' (r-n+1)P(n, r-n+1) =$$
$$(r+1)\sum_{n=0}^{r+1}' P(n, r-n+1) = (r+1)Q(r+1) \tag{16}$$

[1]) Comparing with the general theory in chap. 3 we see that this is a two-dimensional stochastic process with the intensity function $p(n, m, t) = (\lambda+\gamma t)(n+m)$ and the relative transition probability

$$\left\{\begin{matrix} (n'-1, m' | \Pi | n', m'), & (n', m'-1 | \Pi | n', m') \\ (n'-1, m'+2 | \Pi | n', m'), & (n'+2, m'-1 | \Pi | n', m') \end{matrix}\right\} = \left\{\begin{matrix} \frac{\gamma t n'}{(\lambda+\gamma t)(n'+m')}, & \frac{\gamma t m'}{(\lambda+\gamma t)(n'+m')} \\ \frac{\lambda n'}{(\lambda+\gamma t)(n'+m')}, & \frac{\lambda m'}{(\lambda+\gamma t)(n'+m')} \end{matrix}\right\}$$

and all other values equal to 0. We note that Π satisfies the necessary condition

and

$$\sum_{n+m=r} [(n+1)P(n+1, m-2) + (m+1)P(n-2, m+1)] =$$

$$\sum_{n=0}^{r} [(n+1)P(n+1, r-n-2) + (r-n+1)P(n-2, r-n+1)] =$$

$$\sum_{n'=0}^{r-1} n'P(n', r-n'-1) + \sum_{n''=0}^{r-1} (r-n''-1)P(n'', r-n''-1) =$$

$$(r-1) \sum_{n=0}^{r-1} P(n, r-n-1) = (r-1)Q(r-1) . \tag{17}$$

We have thus proved, as we wanted to, that Q satisfies

$$\frac{dQ(r,t)}{dt} = (r+1)\gamma t Q(r+1, t) - r(\lambda + \gamma t)Q(r, t) + (r-1)\lambda Q(r-1, t) \tag{18}$$

which is just (4.7.6).

Multiplying (14) on both sides with $n^k m^l$ and summing over all values of n and m we obtain as a generalization of (4.7.8)

$$\frac{d}{dt}\overline{n^k m^l} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} n^k m^l P'(n, m, t) =$$

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left[((n-1)^k (m+2)^l - n^k m^l)\lambda - (n^k - (n-1)^k) m^l \gamma t\right] nP(n, m, t) +$$

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \left[((m-1)^l (n+2)^k - m^l n^k)\lambda - (m^l - (m-1)^l) n^k \gamma t\right] mP(n, m, t) . \tag{19}$$

In case $k = l = 0$, (19) gives as usual

$$\frac{d}{dt}\overline{n^0 m^0} = \frac{d}{dt} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P(n, m, t) = 0 , \tag{20}$$

i. e., using (15),

$$\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P(n, m, t) = 1 . \tag{21}$$

In case $k = 1$, $l = 0$ and $k = 0$, $l = 1$, respectively, (19) gives

(3.1.6). Furthermore, we see that the process is a *generalized elementary process* (cf. § 3.2) characterized by the operator matrix being *column semi-diagonal* with $l = 2$ (cf. § 7.5). Next, the condition (3.2.6) is satisfied and the operator matrix is thus of type IV (cf. § 7.6). As shown in § 3.3 all the moments then exist and satisfy (3.3.9). Especially we have, as shown in § 3.2, $\overline{n^0 m^0} = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} P(n, m, t) = 1$.

$$\frac{d}{dt}\bar{n} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-\lambda-\gamma t)nP(n, m, t) + \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}2\lambda mP(n, m, t) =$$
$$= -(\lambda+\gamma t)\bar{n} + 2\lambda\bar{m}$$

and (22)

$$\frac{d}{dt}\bar{m} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}2\lambda nP(n, m, t) + \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(-\lambda-\gamma t)mP(n, m, t) =$$
$$= -(\lambda+\gamma t)\bar{m} + 2\lambda\bar{n}$$

which agrees with (1).

In case $k = 2$, $l = 0$, (19) gives

$$\frac{d}{dt}\overline{n^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}[(-2n+1)\lambda-(2n-1)\gamma t]nP(n, m, t) +$$
$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(4n+4)\lambda mP(n, m, t) =$$
$$-2(\lambda+\gamma t)\overline{n^2} + 4\lambda\overline{nm} + (\lambda+\gamma t)\bar{n} + 4\lambda\bar{m} \qquad (\overline{n^2}(0) = 1) \tag{23}$$

in which the initial condition follows from (15).

In case $k = 0$, $l = 2$, (19) gives in the same way

$$\frac{d}{dt}\overline{m^2} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(4m+4)\lambda nP(n, m, t) +$$
$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}[(-2m+1)\lambda-(2m-1)\gamma t]mP(n, m, t) =$$
$$-2(\lambda+\gamma t)\overline{m^2} + 4\lambda\overline{nm} + (\lambda+\gamma t)\bar{m} + 4\lambda\bar{n} \qquad (\overline{m^2}(0) = 0). \tag{24}$$

In case $k = l = 1$, (19) gives finally

$$\frac{d}{dt}\overline{nm} = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\Big[((n-1)(m+2)-nm)\lambda - m\gamma t\Big]nP(n, m, t) +$$
$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\Big[((m-1)(n+2)-mn)\lambda - n\gamma t\Big]mP(n, m, t) =$$
$$-2(\lambda+\gamma t)\overline{nm} + 2\lambda\overline{n^2} + 2\lambda\overline{m^2} - 2\lambda\bar{n} - 2\lambda\bar{m} \qquad (\overline{nm}(0) = 0). \tag{25}$$

As a control we must have $\overline{(n+m)^2}$ satisfying the equation (4.7.12). This is, in fact, the case since, using (23), (24) and (25),

$$\frac{d}{dt}\overline{(n+m)^2} = \frac{d}{dt}(\overline{n^2} + 2\overline{nm} + \overline{m^2}) =$$
$$2(\lambda-\gamma t)(\overline{n^2}+\overline{m^2}) + 4(\lambda-\gamma t)\overline{nm} + (\lambda+\gamma t)(\bar{n}+\bar{m}) =$$
$$2(\lambda-\gamma t)\overline{(n+m)^2} + (\lambda+\gamma t)\overline{(n+m)} \tag{26}$$

which is in agreement with (4.7.12).

In order to solve the three simultaneous differential equations of the first order (23)—(25) we note that (25) may also be written

$$\frac{d}{dt}\overline{nm} = -2(3\lambda+\gamma t)\overline{nm} + 2\lambda\overline{(n+m)^2} - 2\lambda\overline{(n+m)}\,. \tag{27}$$

But in this equation we know both $\overline{n+m}$ and $\overline{(n+m)^2}$ as functions of t since they are given by the expressions found by the discussion of the one-dimensional process, (4.7.2) and (4.7.14) respectively. We see that *the assumption of the process being symmetric in n- and m-particles is essential for this method of solution*. Otherwise, we should have met with a differential equation of the third order which it would be highly unsympathetic to solve! Inserting now (4.7.2) and (4.7.14) into (27) gives

$$\frac{d}{dt}\overline{nm} + 2(3\lambda+\gamma t)\overline{nm} = 4\lambda^2 \exp\left[2\lambda t - \gamma t^2\right]\int_0^t \exp\left[-\lambda t' + \frac{\gamma}{2}t'^2\right]dt'\,. \tag{28}$$

This equation may immediately be integrated and we obtain, due to the initial condition $\overline{nm}(0) = 0$,

$$\overline{nm} = 4\lambda^2 \exp\left[-6\lambda t - \gamma t^2\right]\int_0^t dt' \exp\left[8\lambda t'\right]\int_0^{t'} dt'' \exp\left[-\lambda t'' + \frac{\gamma}{2}t''^2\right]. \tag{29}$$

(29) may be further evaluated as we have

$$\int_0^t dt' \int_0^{t'} dt'' = \int_0^t dt'' \int_{t''}^t dt'\,. \tag{30}$$

(30) into (29) therefore gives

$$\overline{nm} = \frac{\lambda}{2}\exp\left[-6\lambda t - \gamma t^2\right]\cdot$$

$$\left(\exp\left[8\lambda t\right]\int_0^t \exp\left[-\lambda t' + \frac{\gamma}{2}t'^2\right]dt' - \int_0^t \exp\left[7\lambda t' + \frac{\gamma}{2}t'^2\right]dt'\right). \tag{31}$$

We can now evaluate $\overline{n^2}$ and $\overline{m^2}$, as (8) and (31) inserted into (23) and (24), respectively, gives the common equation

$$\frac{d}{dt}\left\{\begin{matrix}\overline{n^2}\\ \overline{m^2}\end{matrix}\right\} + 2(\lambda+\gamma t)\left\{\begin{matrix}\overline{n^2}\\ \overline{m^2}\end{matrix}\right\} = 2\lambda^2 \exp\left[-6\lambda t - \gamma t^2\right]\cdot$$

$$\left(\exp\left[8\lambda t\right]\int_0^t \exp\left[-\lambda t' + \frac{\gamma}{2}t'^2\right]dt' - \int_0^t \exp\left[7\lambda t' + \frac{\gamma}{2}t'^2\right]dt'\right) +$$

$$\frac{\lambda+\gamma t}{2}\left(\exp\left[\lambda t - \frac{\gamma}{2}t^2\right]\left\{\begin{matrix}+\\-\end{matrix}\right\}\exp\left[-3\lambda t - \frac{\gamma}{2}t^2\right]\right) -$$

$$2\lambda\left(\exp\left[\lambda t - \frac{\gamma}{2}t^2\right]\left\{\begin{matrix}-\\+\end{matrix}\right\}\exp\left[-3\lambda t - \frac{\gamma}{2}t^2\right]\right) \tag{32}$$

$$(\overline{n^2}(0) = 1,\ \overline{m^2}(0) = 0)\,.$$

Apart from the different initial conditions and the different signs shown in {}, $\overline{n^2}$ and $\overline{m^2}$ satisfy the *same* differential equation. We can, therefore, solve both equations at the same time. The integration is elementary but tedious and gives, using (8) and (30),

$$\left\{\begin{matrix}\overline{n^2}\\\overline{m^2}\end{matrix}\right\} = \left\{\begin{matrix}\bar{n}\\\bar{m}\end{matrix}\right\} +$$

$$\frac{\lambda}{2}\exp\left[-2\lambda t-\gamma t^2\right]\left(\exp\left[4\lambda t\right]\left\{\begin{matrix}-\\+\end{matrix}\right\}2\right)\int_0^t \exp\left[-\lambda t'+\frac{\gamma}{2}t'^2\right]dt' +$$

$$\frac{\lambda}{2}\exp\left[-6\lambda t-\gamma t^2\right]\int_0^t \exp\left[7\lambda t'+\frac{\gamma}{2}t'^2\right]dt' . \tag{33}$$

Comparing the fluctuation formula (33) with the corresponding formula for the one-dimensional model, (4.7.14), it is not immediately evident that the former formula gives a larger fluctuation than the latter—apart from the initial point since they are both subjected to the same initial condition $\overline{n^2}(0) = 1$. For larger values of λt we have, however, approximately $\bar{n} \sim \bar{m} \sim \frac{1}{2}\exp\left[\lambda t - \frac{\gamma}{2}t^2\right]$ and thus approximately

$$\overline{n^2} \sim \overline{m^2} \sim \bar{n} + 2\lambda\bar{n}^2\int_0^t \exp\left[-\lambda t'+\frac{\gamma}{2}t'^2\right]dt' +$$

$$\frac{\lambda}{2}\exp\left[-6\lambda t-\gamma t^2\right]\int_0^t \exp\left[7\lambda t'+\frac{\gamma}{2}t'^2\right]dt' . \tag{34}$$

Comparing (34) with (4.7.14) it is seen that for the same value of $\bar{n}$ the two first terms are identical. Since the third term of (34) is *positive*, the fluctuation is, therefore, in fact increased by splitting the one-dimensional process into a two-dimensional process, in agreement with an observation on page 116.

As a control of the correctness of the important formula (33) we have

$$\overline{n^2} + 2\overline{nm} + \overline{m^2} = \overline{(n+m)^2} =$$

$$\exp\left[\lambda t - \frac{\gamma}{2}t^2\right] + 2\lambda\exp\left[2\lambda t-\gamma t^2\right]\int_0^t \exp\left[-\lambda t'+\frac{\gamma}{2}t'^2\right]dt' \tag{35}$$

as given by (4.7.14). Inserting (31) and (33) we find that this relation is in fact fulfilled.

For the numerical evaluation of (33) we can use the function $\varphi(x)$ given by (4.7.15) and tabulated in chap. 8. We thus have

$$\begin{Bmatrix} \overline{n^2} \\ \overline{m^2} \end{Bmatrix} = \begin{Bmatrix} \bar{n} \\ \bar{m} \end{Bmatrix} + \exp\left[-2\lambda t - \gamma t^2\right] \cdot$$

$$\left(\exp\left[4\lambda t\right] \begin{Bmatrix} - \\ + \end{Bmatrix} 2\right) \exp\left[-\frac{\lambda^2}{2\gamma}\right] \frac{\lambda}{\sqrt{2\gamma}} \left(\psi\left(\frac{\gamma t - \lambda}{\sqrt{2\gamma}}\right) + \psi\left(\frac{\lambda}{\sqrt{2\gamma}}\right)\right) +$$

$$\exp\left[-6\lambda t - \gamma t^2\right] \exp\left[-\frac{49\lambda^2}{2\gamma}\right] \frac{\lambda}{\sqrt{2\gamma}} \left(\psi\left(\frac{\gamma t + 7\lambda}{\sqrt{2\gamma}}\right) - \psi\left(\frac{7\lambda}{\sqrt{2\gamma}}\right)\right). \quad (36)$$

Of this formula we shall make extensive use in chap. 6.

§ 4.9. Having calculated the first and second moments of our two-dimensional stochastic process we shall now investigate the numerical values of the probabilities $P(n, m, t)$ themselves, or more correctly of the *marginal distributions* in which we are interested

$$P(n, t) = \sum_{m=0}^{\infty} P(n, m, t)$$

and (1)

$$P(m, t) = \sum_{n=0}^{\infty} P(n, m, t).$$

The functions $P(n, m, t)$ are solutions of the double-infinite system of simultaneous differential equations (4.8.14). As previously discussed (cf. page 112) the corresponding one-dimensional process is already so complicated that it seems impossible to obtain the solutions in a closed form. As discussed at the same place, even if this were possible, the solutions would certainly be so complicated functions that they would be useless for the practical problems to which we shall apply them later (cf. chap. 6). It is, however, *a natural idea that the Pólya distribution will represent a sufficiently good approximation to our marginal distributions if for each value of the 'time' t we adjust the two Pólya parameters to give the same values of the first and second moments as the model.* The Pólya distribution belongs namely, as we have seen in § 4.4, to the same family of distributions, the Furry process being a special case of the Pólya process and our process being a generalization of the Furry process.

In order to test this hypothesis before we make use of it in the applications to the theory of cosmic radiation (cf. chap. 6) we must, however, obtain some idea, if only a rough one, about the numerical values of the functions $P(n, m, t)$ and thus of the marginal distributions. As previously discussed we must have recourse to the *product-*

integral, given in (3.1.17), for the numerical evaluation of the solutions. Here we have two possibilities. Either we may include many terms in the iteration series, being the two-dimensional generalization of (2.4.1), and then repeat the exponentiable continuation (cf. § 2.7) a few times only[1]), or we may include a few terms only and then repeat the continuation many times. As the evaluation of products of two-dimensional matrices is rather unhandy in practice we have preferred the latter procedure[2]), choosing the interval $\Delta t = t - s$ so small that we need only keep the *linear* terms in each of the iteration series. *This procedure means a direct application of the definition of the product-integral*, (3.1.17), approximating it by the expression $\underset{i}{\Pi}(\mathbf{1} + \boldsymbol{A}(t_i)\Delta_i)$ (cf. § 2.5). In other words, the idea is simply to expand all the functions in their Taylor series but keep only the linear terms. Thus, assuming the functions to follow their tangents approximately in a small time interval Δt we can calculate, by means of the equations (4.8.13), the values of $P(n, m, t)$ at the time $t + \Delta t$ when the values are known at the time t. We can then repeat the calculations and obtain the values at the time $t + 2\Delta t$, and so on. It is, firstly, clear that the solutions obtained in this way can only be approximately correct, unless Δt is taken very small which would make the calculations extremely lengthy. We do not need, however, any higher accuracy than say of the order 10-20% and since the error will, presumably, be of the order of the first term neglected, i. e. $(\Delta t)^2$, we can choose $\Delta t = 0{\cdot}1$. Next, it is clear that this is only so for $P(n, m)$-functions not too far out in the tail, the accuracy growing worse and worse the higher the values of n and m, because all the coefficients in (4.8.13) increase with increasing values of n and m. Finally, it is clear in advance that this calculation cannot be pushed too far since from a certain step the term $1 - (n + m)(\lambda + \gamma t)\Delta t$ will be *negative* and sooner or later we shall thus

[1]) We note that as shown in example 1 § 2.7 this exponentiable continuation may be carried through to arbitrary high values of t, as the present intensity function is actually of the form to which the example applies.

[2]) We observe that we would also have to repeat the continuation several times in the former procedure due to the condition limiting the convergence of the iteration series, viz. $t - s < \frac{1}{2t(\lambda + \gamma t)} = \frac{1}{4\left(1 + \frac{t}{4}\right)} < \frac{1}{4}$, as $t = 2$ (cf. the example quoted above and the following discussion of the numerical values of λ and γ).

obtain *negative* values for some $P(n, m)$-functions despite these functions being *probabilities*, i.e. *positive*. A control of the accuracy of and extent to which we may carry the calculations is, however, offered by the fact that we know the *exact* values of the first and second moments, which are given by (4.8.8) and (4.8.33).

Before we proceed with the calculations we note that although our distribution depends formally on *two* parameters λ and γ, being measures of the 'birth' and 'death' rates, respectively, it really depends on only one, namely their ratio. Dividing (4.8.14) by λ on both sides gives

$$\frac{d}{d(\lambda t)} P(n, m, t) = (n+1)\frac{\gamma}{\lambda^2}\lambda t P(n+1, m, t) + (m+1)\frac{\gamma}{\lambda^2}\lambda t P(n, m+1, t) -$$

$$(n+m)\left(1+\frac{\gamma}{\lambda^2}\lambda t\right) P(n, m, t) +$$

$$(n+1)P(n+1, m-2, t) + (m+1)P(n-2, m+1, t) \tag{2}$$

which shows that if we introduce a new time scale $\lambda t \to t$ and make the substitution $\frac{\gamma}{\lambda^2} \to \gamma$ we just obtain the same system of differential equations (4.8.14) with $\lambda = 1$. We shall now evaluate the solutions of these equations numerically, approximating them with solutions of the equations (4.8.13) with $\lambda = 1$, i.e. we have to solve the equations

$$P(n, m, t+\Delta t) = (n+1)\gamma t\Delta t P(n+1, m, t) + (m+1)\gamma t\Delta t P(n, m+1, t) +$$

$$(1-(n+m)(1+\gamma t)\Delta t)P(n, m, t) +$$

$$(n+1)\Delta t P(n+1, m-2, t) + (m+1)\Delta t P(n-2, m+1, t) \tag{3}$$

$$(P(n, m, 0) = \delta_{n,1}\cdot\delta_{m,0}).$$

As a valuable check on the computation we have

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P(n, m, t) = 1. \tag{4}$$

This follows from the fact that applying the operator in (3) does not change the sum and since this sum equals one for $t = 0$ it will remain one. (3) can, namely, in our matrix symbolism (cf. chap. 3) be written

$$P(n, m, t+\Delta t) = \sum_{n'=0}^{\infty}\sum_{m'=0}^{\infty}(n, m|\boldsymbol{T}|n', m')P(n', m', t) \tag{5}$$

in which $(n, m|\boldsymbol{T}|n', m')$ is the matrix element of the transition probability operator, (cf. (2.7.4)), giving the asymptotic value of the transition probability *from* the state (n', m') *to* the state (n, m). We therefore have, due to (2.7.4) and (3.1.13),

Table 1. *The probabilities $P(n, m, t+\Delta t)$ for $t+\Delta t = 0{\cdot}5$.*

$\gamma = 0{\cdot}25$ $t = 0{\cdot}4$ $\Delta t = 0{\cdot}1$ $\gamma t = 0{\cdot}1$ $(1+\gamma t)\Delta t = 0{\cdot}11$

n \ m	0	1	2	3	4	5	6	$\sum_m$	$n\sum_m$	$n^2\sum_m$
0	·01823	·00939	·25237	·00024	·00003	·00136	0	·28162	0	0
1	·57428	·00219	·00057	·02165	0	0	0	·59869	·59869	·59869
2	·00181	·09854	·00010	0	·00080	0	0	·10125	·20250	·40500
3	·00038	·00006	·00680	0	0	0	0	·00724	·02172	·06516
4	·01082	0	0	0	0	0	0	·01082	·04328	·17312
5	0	·00040	0	0	0	0	0	·00040	·00200	·01000
$\sum_n$	·60552	·11058	·25984	·02189	·00083	·00136	0	1·00002	·86819	1·25197
$m\sum_n$	0	·11058	·51968	·06567	·00332	·00680	0	·70605		
$m^2\sum_n$	0	·11058	1·03936	·19701	·01328	·03400	0	1·39423		

$\overline{n} = 0{\cdot}86819$ $(\overline{n}_{exact} = 0{\cdot}9071)$ $\overline{n^2} = 1{\cdot}25197$ $(\overline{n^2}_{exact} = 1{\cdot}386)$

$\overline{m} = 0{\cdot}70605$ $(\overline{m}_{exact} = 0{\cdot}6909)$ $\overline{m^2} = 1{\cdot}39423$ $(\overline{m^2}_{exact} = 1{\cdot}445)$

$$\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} P(n, m, t+\Delta t) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n'=0}^{\infty}\sum_{m'=0}^{\infty}(n, m|\boldsymbol{T}|n', m')P(n', m', t) =$$

$$\sum_{n'=0}^{\infty}\sum_{m'=0}^{\infty}\left[\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}(n, m|\boldsymbol{T}|n', m')\right]P(n', m', t) =$$

$$\sum_{n'=0}^{\infty}\sum_{m'=0}^{\infty} 1 \cdot P(n', m', t) = 1\,. \qquad (6)$$

The interchanging of the order of summation is here legitimate because all the sums are in reality *finite* sums since we start with a finite number of non-vanishing $P(n, m)$ functions, due to the initial condition, and our probability operator $\boldsymbol{T}$ has only a finite number of non-vanishing terms for each fixed n', m' (cf. [1]) p. 117).

In order to keep down the rounding-off errors and thereby fully utilizing the check (4) we have performed the computation to five places in spite of the fact that presumably only the two first places are accurate. Having calculated all non-vanishing $P(n, m, t+\Delta t)$-functions and checked that their sum is really one to five places, we next calculate the marginal distributions (1) thus obtaining

Table 2. *The probabil*

$\gamma = 0{\cdot}25$				$t = 0{\cdot}9$		$\gamma t = 0{\cdot}22$
n \ m	0	1	2	3	4	5
0	·05882	·03923	·19972	·00474	·00246	·0077
1	·30913	·01594	·01128	·05217	·00219	·0008
2	·01593	·12668	·00762	·00387	·01790	·0005
3	·00683	·00501	·03898	·00202	·00086	·0045
4	·02605	·00239	·00159	·01058	·00032	·0000
5	·00065	·00895	·00062	·00027	·00223	0
6	·00026	·00024	·00272	·00005	0	·0002
7	·00081	·00007	·00003	·00049	0	0
8	0	·00031	0	0	0	0
9	0	0	·00004	0	0	0
10	·00001	0	0	0	0	0
$\sum_n$	·41849	·10882	·26260	·07419	·02596	·0140
$m\sum_n$	0	·10882	·52520	·22257	·10384	·0701
$m^2\sum_n$	0	·10882	1·05040	·66771	·41536	·3507

$\bar{n} = 1{\cdot}18165$ $(\bar{n}_{exact} = 1{\cdot}2214)$ $\overline{n^2} = 2{\cdot}81915$ $(\overline{n^2}_{exact} = 3{\cdot}246)$

$\bar{m} = 1{\cdot}15810$ $(\bar{m}_{exact} = 1{\cdot}1775)$ $\overline{m^2} = 2{\cdot}92417$ $(\overline{m^2}_{exact} = 3{\cdot}344)$

an approximation to the probability at the time $t + \Delta t$ of finding n n-particles independent of the number of m-particles present and vice versa for $P(m, t + \Delta t)$.

Since, in fact, we are only interested in the marginal distributions it would be much easier to obtain differential equations for these distributions and solving them direct. This is, unfortunately, impossible. From (2) with $\lambda = 1$ we obtain, namely,

$$\frac{d}{dt} P(n, t) = (n + 1)(1 + \gamma t) P(n + 1, t) - n(1 + \gamma t) P(n, t) -$$

$$\sum_{m=0}^{\infty} m P(n, m, t) + \sum_{m=0}^{\infty} m P(n - 2, m, t), \tag{7}$$

and the analogue for $P(m, t)$, and these equations contain two terms which cannot be expressed by means of the marginal distributions.

$, m, t+\Delta t)$ *for* $t+\Delta t=1{\cdot}0$.

$\Delta t = 0{\cdot}1$	$(1+\gamma t)\Delta t = 0{\cdot}1225$					
6	7	8	9	$\sum_m$	$n\sum_m$	$n^2\sum_m$
·00021	·00007	·00024	0	·31328	0	0
·00326	·00006	·00005	·00008	·39504	·39504	·39504
·00018	·00088	0	0	·17362	·34724	·69448
·00005	0	·00010	0	·05835	·17505	·52515
·00075	0	0	0	·04176	·16704	·66816
0	0	0	0	·01272	·06360	·31800
0	0	0	0	·00349	·02094	·12564
0	0	0	0	·00140	·00980	·06860
0	0	0	0	·00031	·00248	·01984
0	0	0	0	·00004	·00036	·00324
0	0	0	0	·00001	·00010	·00100
·00445	·00101	·00039	·00008	1·00002	1·18165	2·81915
·02670	·00707	·00312	·00072	1·15819		
·16020	·04949	·02496	·00648	2·92417		

Finally, we calculate the moments of first and second order

$$
\begin{aligned}
\bar{n} &= \sum_n nP(n, t+\Delta t)\\
\bar{m} &= \sum_m mP(m, t+\Delta t)\\
\overline{n^2} &= \sum_n n^2P(n, t+\Delta t)\\
\overline{m^2} &= \sum_m m^2P(m, t+\Delta t)
\end{aligned}
\tag{8}
$$

and compare them with the exact values obtained from (4.8.8) and (4.8.33), respectively, (denoted by $\bar{n}_{exact}$, $\overline{n^2}_{exact}$ etc. in tables 1–5).

For the numerical value of γ we have chosen

$$\gamma = \frac{1}{4} \tag{9}$$

Table 3. *The probab*

	$\gamma = 0{\cdot}25$				$t = 1{\cdot}4$			$\gamma t = 0$
n \ m	0	1	2	3	4	5	6	
0	·10186	·05897	·11494	·01077	·00671	·00726	·00119	·0
1	·16237	·02665	·02570	·04074	·00823	·00451	·00491	·0
2	·03151	·08383	·02072	·01420	·02084	·00490	·00273	·0
3	·01362	·01429	·03645	·01063	·00727	·00939	·00251	·0
4	·01994	·00856	·00861	·01604	·00529	·00346	·00413	·0
5	·00261	·01042	·00510	·00441	·00700	·00236	·00145	·0
6	·00133	·00213	·00567	·00259	·00206	·00288	·00100	·0
7	·00122	·00108	·00131	·00273	·00111	·00088	·00099	·0
8	·00021	·00104	·00068	·00084	·00122	·00033	·00048	·0
9	·00010	·00012	·00057	·00035	·00017	·00054	0	[illegible]
10	·00008	·00009	·00014	·00024	·00021	0	·00048	[illegible]
11	·00003	·00007	·00003	·00008	0	0	0	[illegible]
12	0	·00001	·00005	0	0	0	0	[illegible]
13	0	·00001	0	·00006	0	0	0	[illegible]
$\sum_n$	·33488	·20727	·21907	·10368	·06011	·03651	·01987	[illegible]
$m\sum_n$	0	·20727	·43904	·31104	·24044	·18255	·11922	[illegible]
$m^2\sum_n$	0	·20727	·87988	·93312	·96176	·91275	·71532	[illegible]

$\bar{n} = 1{\cdot}64283$ $(\bar{n}_{exact} = 1{\cdot}695)$ $\overline{n^2} = 5{\cdot}68659$ $(\overline{n^2}_{exact} = 6{\cdot}636)$

$\bar{m} = 1{\cdot}63821$ $(\bar{m}_{exact} = 1{\cdot}687)$ $\overline{m^2} = 5{\cdot}70091$ $(\overline{m^2}_{exact} = 6{\cdot}675)$

which value, due to (4.8.10), corresponds to

$$\bar{n}_{max} = \bar{m}_{max} = \frac{1}{2} e^2 = 3{\cdot}695 \sim 4 \quad \text{for} \quad t_{max} = 4\,, \tag{10}$$

i. e. a cascade shower of medium size. The results are shown in tables 1–3 giving the schemes for $t+\Delta t = 0{\cdot}5$, $1{\cdot}0$ and $1{\cdot}5$. In the next scheme, $t+\Delta t = 1{\cdot}6$, negative values appeared for the first time and we therefore decided to carry on the computation with a smaller value of Δt, namely $\Delta t = 0{\cdot}05$. Nevertheless, negative values soon appeared again but, of course, of smaller magnitudes. The computation was performed up to $t+\Delta t = 2{\cdot}0$ and then to four places up to $t+\Delta t = 2{\cdot}10$. From this point on the negative values, already amounting to a maximum value of $-0{\cdot}0197$ (for $n = 11$

, $m, t+\Delta t)$ *for* $t+\Delta t = 1{\cdot}5$.

$\Delta t = 0{\cdot}1$ $(1+\gamma t)\Delta t = 0{\cdot}135$

8	9	10	11	12	13	$\sum_m$	$n\sum_m$	$n^2\sum_m$
049	·00009	·00004	·00003	·00001	0	·30292	0	0
049	·00042	·00008	·00003	·00002	·00001	·27517	·27517	·27517
062	·00031	·00027	·00004	·00003	0	·18277	·36554	·73108
133	·00033	·00011	·00013	0	0	·09745	·29235	·87705
059	·00050	·00016	0	·00012	0	·06849	·27396	1·09584
034	·00034	·00010	0	0	0	·03586	·17930	·89650
060	0	0	0	0	0	·01869	·11214	·67284
	·00056	0	0	0	0	·01038	·07266	·50862
	0	0	0	0	0	·00484	·03872	·30976
	0	0	0	0	0	·00185	·01665	·14985
	0	0	0	0	0	·00124	·01240	·12400
	0	0	0	0	0	·00021	·00231	·02541
	0	0	0	0	0	·00006	·00072	·00864
	0	0	0	0	0	·00007	·00091	·01183
452	·00255	·00076	·00023	·00018	·00001	1·00000	1·64283	5·68659
616	·02295	·00760	·00253	·00216	·00013	1·63821		
928	·20655	·07600	·02783	·02592	·00169	5·70091		

and $m = 12$), were considered to become too dangerous and our approximation was, therefore, thought to give too distorted a picture to allow the computations to be pushed further. The results are given in tables 4–5 giving the schemes for $t+\Delta t = 2{\cdot}0$ and $2{\cdot}10$. It is seen that in all the tables the first moments agree within 5% and the quadratic ones within 10%, the agreement thus being adequate.

We can now finally test our hypothesis of the Pólya distribution being a sufficiently good approximation for the practical purpose to which we are going to apply our stochastic process (cf. chap. 6). For each value of t we adjust the two parameters of the Pólya distribution (4.4.1) by means of (4.4.9) and (4.4.10), respectively, to give the same values of the first and second moments as found by the numerical computation of the marginal distributions. The results are shown for $P(m, t)$ in tables 6–10 and figs. 5–8 and it is

Table 4. *The probabilitie*

$\gamma = 0{\cdot}25$ $t = 1{\cdot}95$ $\gamma t = 0{\cdot}4875$

$n \backslash m$	0	1	2	3	4	5	6	7	8	9	10	11
0	·14188	·05978	·06466	·01272	·00808	·00512	·00172	·00084	·00047	·00018	·00008	·00004
1	·09285	·02792	·03040	·02682	·01072	·00628	·00428	·00180	·00096	·00054	·00024	·00012
2	·03600	·05041	·02464	·01855	·01634	·00772	·00473	·00313	·00151	·00083	·00047	·00022
3	·01430	·01787	·02587	·01495	·01138	·00933	·00510	·00320	·00210	·00109	·00061	·00037
4	·01255	·01085	·01240	·01410	·00926	·00695	·00544	·00321	·00205	·00130	·00075	·00034
5	·00357	·00816	·00780	·00782	·00798	·00556	·00415	·00315	·00191	·00128	·00067	·00067
6	·00178	·00340	·00560	·00513	·00485	·00454	·00327	·00238	·00185	·00093	·00107	−·00018
7	·00104	·00183	·00270	·00357	·00322	·00291	·00254	·00194	·00113	·00146	−·00030	·00183
8	·00038	·00110	·00152	·00190	·00220	·00192	·00175	·00119	·00156	−·00036	·00226	−·00224
9	·00018	·00048	·00091	·00110	·00122	·00135	·00093	·00165	−·00036	·00270	−·00265	·00429
10	·00009	·00024	·00044	·00065	·00076	·00064	·00117	−·00038	·00238	−·00270	·00420	−·00437
11	·00004	·00013	·00023	·00034	·00037	·00068	−·00022	·00180	−·00231	·00394	−·00416	·00479
12	·00002	·00005	·00012	·00015	·00034	−·00013	·00118	−·00170	·00352	−·00386	·00449	−·00306
13	·00001	·00003	·00005	·00014	−·00004	·00067	−·00102	·00216	−·00294	·00264	−·00270	·00097
14	0	·00001	·00004	−·00002	·00027	−·00055	·00128	−·00181	·00270	−·00202	·00276	0
15	0	0	0	·00010	−·00021	·00070	−·00103	·00162	−·00145	·00062	0	·00045
16	0	0	·00002	−·00006	·00024	−·00050	·00062	−·00084	·00037	−·00102	0	0
17	0	0	−·00002	·00007	−·00018	·00038	−·00039	·00071	0	0	0	0
18	0	0	·00002	−·00005	·00012	−·00017	·00009	0	·00010	0	0	0
19	0	0	−·00001	·00002	−·00005	·00003	−·00011	0	0	0	0	0
20	0	0	0	−·00001	·00004	0	0	0	0	0	0	0
$\sum_n$	·30469	·18226	·17739	·10799	·07691	·05343	·03548	·02405	·01555	·00755	·00779	·00424
$m\sum_n$	0	·18226	·35478	·32397	·30764	·26715	·21288	·16835	·12440	·06795	·07790	·0466
$m^2\sum_n$	0	·18226	·70956	·97191	1·23056	1·33575	1·27728	1·17845	·99520	·61155	·77900	·5130

$\bar{n} = 2{\cdot}15233$ ($\bar{n}_{exact} = 2{\cdot}242$) $\overline{n^2} = 10{\cdot}26931$ ($\overline{n^2}_{exact} = 12{\cdot}14$)

$\bar{m} = 2{\cdot}17132$ ($\bar{m}_{exact} = 2{\cdot}240$) $\overline{m^2} = 10{\cdot}30820$ ($\overline{m^2}_{exact} = 12{\cdot}15$)

seen that the agreement is, of course, only rough but on the other side, bearing in mind that our process is in itself only an approximation to the real multiplication process, sufficiently good to allow us to assume that results obtained by using this Pólya approximation will give the same trend as would the exact distribution.

$P(n, m, t+\Delta t)$ for $t+\Delta t = 2\cdot0$.

$\Delta t = 0\cdot05$ $(1+\gamma t)\Delta t = 0\cdot074375$

12	13	14	15	16	17	18	19	20	21	Σ_m	$n\Sigma_m$	$n^2\Sigma_m$
·00001	0	0	0	0	0	0	0	0	0	·29558	0	0
·00006	·00003	·00001	·00001	0	0	0	0	0	0	·20304	·20304	·20304
·00011	·00005	·00004	0	·00003	−·00001	·00001	0	0	0	·16478	·32956	·65912
·00015	·00013	−·00001	·00009	−·00006	·00007	−·00004	·00003	−·00001	0	·10652	·31956	·95868
·00034	−·00005	·00028	−·00019	·00022	−·00015	·00004	−·00004	·00001	−·00001	·07965	·31860	1·27440
·00011	·00061	−·00049	·00056	−·00041	·00031	−·00015	·00011	0	0	·05315	·26575	1·32875
·00117	−·00098	·00131	−·00096	·00079	−·00037	·00009	0	·00002	0	·03569	·21414	1·28484
−·00163	·00209	−·00177	·00109	−·00073	·00021	−·00028	0	0	0	·02285	·15995	1·11965
·00304	−·00263	·00247	−·00135	·00118	0	0	0	0	0	·01589	·12712	1·01696
·00376	·00361	−·00198	·00058	0	·00020	0	0	0	0	·01045	·09405	·84645
·00314	−·00256	·00081	−·00139	0	0	0	0	0	0	·00312	·03120	·31200
·00298	·00319	0	0	0	0	0	0	0	0	·00584	·06424	·70664
·00103	0	·00052	0	0	0	0	0	0	0	·00267	·03204	·38448
·00208	0	0	0	0	0	0	0	0	0	−·00211	−·02743	−·35659
0	0	0	0	0	0	0	0	0	0	·00266	·03724	·52136
0	0	0	0	0	0	0	0	0	0	·00080	·01200	·18000
0	0	0	0	0	0	0	0	0	0	−·00117	−·01872	−·29952
0	0	0	0	0	0	0	0	0	0	·00057	·00969	·16473
0	0	0	0	0	0	0	0	0	0	·00011	·00198	·03564
0	0	0	0	0	0	0	0	0	0	−·00012	−·00228	−·04332
0	0	0	0	0	0	0	0	0	0	·00003	·00060	·01200
−·00151	·00349	·00119	−·00156	·00102	·00026	−·00033	·00010	·00002	−·00001	1·00000	2·15233	10·26931
·01812	·04537	·01666	−·02340	·01632	·00442	−·00594	·00190	·00040	−·00021	2·17132		
−·21744	·58981	·23324	−·35100	·26112	·07514	−·10692	·03610	·00800	−·00441	10·30820		
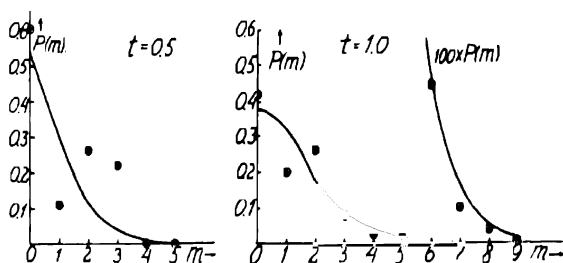
FIG. 5-6. *The marginal probabilities P(m, t) of the shower model, given by ⊕, compared with the Pólya approximation, given by the full curve, for t = 0·5 and 1·0.*

9*

Table 5. *The probabilities*

$\gamma = 0{\cdot}25$ $\quad t = 2{\cdot}05$ $\quad \gamma t = 0{\cdot}5125$

n \ m	0	1	2	3	4	5	6	7	8	9	10	11
0	·1495	·0586	·0582	·0122	·0079	·0046	·0017	·0008	·0005	·0002	·0001	0
1	·0843	·0276	·0303	·0249	·0108	·0063	·0041	·0019	·0010	·0005	·0002	·0001
2	·0358	·0460	·0246	·0186	·0156	·0079	·0049	·0032	·0016	·0009	·0005	·0002
3	·0141	·0179	·0243	·0152	·0116	·0092	·0053	·0034	·0022	·0012	·0006	·0003
4	·0115	·0110	·0126	·0137	·0096	·0073	·0056	·0035	·0023	·0014	·0009	·0004
5	·0036	·0078	·0080	·0081	·0081	·0059	·0045	·0034	·0022	·0015	·0007	·0010
6	·0019	·0038	·0056	·0054	·0052	·0048	·0036	·0027	·0021	·0010	·0016	−·000[illegible]
7	·0010	·0019	·0029	·0037	·0035	·0032	·0028	·0023	·0012	·0022	−·0013	·0041
8	·0004	·0011	·0016	·0021	·0024	·0022	·0021	·0017	·0026	−·0016	·0053	−·007[illegible]
9	·0002	·0005	·0010	·0013	·0014	·0016	·0010	·0022	−·0016	·0058	−·0082	·012[illegible]
10	·0001	·0003	·0005	·0008	·0009	·0006	·0017	−·0013	·0054	−·0081	·0139	−·01[illegible]
11	0	·0001	·0002	·0004	·0004	·0010	−·0009	·0043	−·0071	·0129	−·0180	·020[illegible]
12	0	0	·0001	·0001	·0005	−·0006	·0028	−·0052	·0104	−·0147	·0190	−·010[illegible]
13	0	0	0	·0002	−·0002	·0015	−·0030	·0072	−·0119	·0159	−·0177	·0151
14	0	0	0	−·0001	·0007	−·0016	·0042	−·0077	·0110	−·0135	·0106	−·0071
15	0	0	0	·0002	−·0007	·0020	−·0040	·0069	−·0077	·0096	−·0056	·005[illegible]
16	0	0	0	−·0002	·0008	−·0018	·0037	−·0054	·0052	−·0062	·0011	0
17	0	0	0	·0002	−·0007	·0015	−·0026	·0028	−·0024	·0008	−·0022	0
18	0	0	0	−·0001	·0005	−·0009	·0017	−·0012	·0023	0	0	0
19	0	0	0	·0001	−·0004	·0006	−·0009	·0002	0	0	0	0
20	0	0	0	0	·0002	−·0002	·0001	−·0002	0	0	0	0
21	0	0	0	0	0	·0001	0	0	0	0	0	0
Σ_n	·3024	·1766	·1699	·1068	·0781	·0552	·0384	·0255	·0193	·0098	·0015	·014[illegible]
$m\Sigma_n$	0	·1766	·3398	·3204	·3124	·2760	·2304	·1785	·1544	·0882	·0150	·155[illegible]
$m^2\Sigma_n$	0	·1766	·6796	·9612	1·2496	1·3800	1·3824	1·2495	1·2352	·7938	·1500	1·70[illegible]

$\bar{n} = 2{\cdot}2862 \quad (\bar{n}_{exact} = 2{\cdot}353)$ $\qquad \overline{n^2} = 11{\cdot}4092 \quad (\overline{n^2}_{exact} = 13{\cdot}44)$

$\bar{m} = 2{\cdot}2880 \quad (\bar{m}_{exact} = 2{\cdot}352)$ $\qquad \overline{m^2} = 11{\cdot}6456 \quad (\overline{m^2}_{exact} = 13{\cdot}45)$
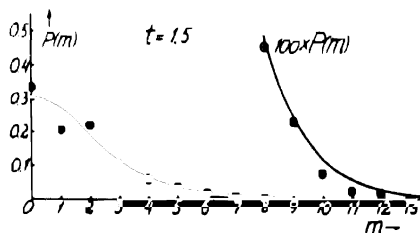
FIG. 7. *The marginal probabilities P(m, t) of the shower model, given by ⊕, compared with the Pólya approximation, given by the full curve, for t = 1·5.*

$P(n, m, t+\Delta t)$ *for* $t+\Delta t = 2{\cdot}10$.

$\Delta t = 0{\cdot}05$ $\qquad (1+\gamma t)\Delta t = 0{\cdot}075625$

12	13	14	15	16	17	18	19	20	21	Σ_m	$n\Sigma_m$	$n^2\Sigma_m$
0	0	0	0	0	0	0	0	0	0	·2943	0	0
0	0	0	0	0	0	0	0	0	0	·1920	·1920	·1920
·0001	0	0	0	0	0	0	0	0	0	·1599	·3198	·6396
·0001	·0002	−·0001	·0002	−·0001	·0002	−·0001	·0001	0	0	·1058	·3174	·9522
·0004	−·0002	·0006	−·0006	·0008	−·0006	·0006	−·0002	·0001	0	·0807	·3228	1·2912
·0005	·0014	−·0015	·0018	−·0017	·0014	−·0014	·0004	−·0001	0	·0546	·2730	1·3650
·0026	·0029	·0039	−·0035	·0032	−·0021	·0012	−·0004	·0004	0	·0393	·2358	1·4148
·0048	·0068	−·0068	·0065	−·0049	·0032	−·0017	·0002	0	0	·0260	·1820	1·2740
·0102	−·0114	·0107	−·0079	·0047	−·0023	·0005	−·0007	0	0	·0167	·1336	1·0688
·0123	·0150	−·0109	·0089	−·0038	·0037	0	0	0	0	·0185	·1665	1·4985
·0169	·0172	·0124	−·0083	·0012	0	·0004	0	0	0	·0035	·0350	·3500
·0197	·0124	−·0069	·0016	−·0032	0	0	0	0	0	−·0024	−·0264	−·2904
·0162	−·0079	·0095	0	0	0	0	0	0	0	·0137	·1644	1·9728
·0124	·0019	0	·0009	0	0	0	0	0	0	−·0025	−·0325	−·4225
·0010	·0046	0	0	0	0	0	0	0	0	−·0062	−·0868	−1·2152
0	0	0	0	0	0	0	0	0	0	·0089	·1335	2·0025
·0007	0	0	0	0	0	0	0	0	0	−·0021	−·0336	−·5376
0	0	0	0	0	0	0	0	0	0	−·0026	−·0442	−·7514
0	0	0	0	0	0	0	0	0	0	·0023	·0414	·7452
0	0	0	0	0	0	0	0	0	0	−·0004	−·0076	−·1444
0	0	0	0	0	0	0	0	0	0	−·0001	−·0020	−·0400
0	0	0	0	0	0	0	0	0	0	·0001	·0021	·0441
·0006	−·0065	·0109	−·0004	−·0038	·0035	−·0005	−·0006	·0004	0	1·0000	2·2862	11·4092
−·0072	−·0845	·1526	−·0060	−·0608	·0595	−·0090	−·0114	·0080	0	2·2880		
−·0864	−1·0985	2·1364	−·0900	−·9728	1·0115	−·1620	−·2166	·1600	0	11·6456		
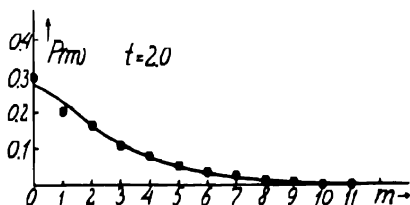
Fig. 8. *The marginal probabilities P(m, t) of the shower model, given by ⊕, compared with the Pólya approximation, given by the full curve, for t = 2·0.*

Table 6–10. *The probabilities $P(m, t)$ compared with the Pólya approximation.*

Table 6. $t = 0{\cdot}5$.

m	Model	Pólya
0	0·60552	0·53503
1	0·11058	0·29775
2	0·25984	0·11438
3	0·02189	0·03738
4	0·00083	0·01113
5	0·00136	0·00312
≥ 6	0	0·00121
Σ	1·00002	1·00000

Table 7. $t = 1{\cdot}0$.

m	Model	Pólya
0	0·41849	0·37277
1	0·19882	0·31592
2	0·26260	0·17624
3	0·07419	0·08130
4	0·02596	0·03358
5	0·01403	0·01290
6	0·00445	0·00471
7	0·00101	0·00165
8	0·00039	0·00056
9	0·00008	0·00019
≥ 10	0	0·00018
Σ	1·00002	1·00000

Table 8. $t = 1{\cdot}5$.

m	Model	Pólya
0	0·33488	0·30465
1	0·20727	0·27099
2	0·21997	0·18245
3	0·10368	0·10969
4	0·06011	0·06199
5	0·03651	0·03369
6	0·01987	0·01783
7	0·00946	0·00925
8	0·00452	0·00473
9	0·00255	0·00239
10	0·00076	0·00119
11	0·00023	0·00059
12	0·00018	0·00029
13	0·00001	0·00014
≥ 14	0	0·00013
Σ	1·00000	1·00000

Table 9. $t = 2{\cdot}0$.

m	Model	Pólya
0	0·30469	0·27195
1	0·18226	0·22877
2	0·17739	0·16629
3	0·10799	0·11454
4	0·07691	0·07671
5	0·05343	0·05050
6	0·03548	0·03286
7	0·02405	0·02120
8	0·01555	0·01359
9	0·00755	0·00867
10	0·00779	0·00551
11	0·00424	0·00349
12	− 0·00151	0·00220
13	0·00349	0·00139
14	0·00119	0·00087
15	− 0·00156	0·00055
16	0·00102	0·00034
17	0·00026	0·00021
18	− 0·00033	0·00013
19	0·00010	0·00008
20	0·00002	0·00005
21	− 0·00001	0·00003
≥ 22	0	0·00007
Σ	1·00000	1·00000

Table 10. $t = 2{\cdot}10$.

m	Model	Pólya
0	0·3024	0·27033
1	0·1766	0·22072
2	0·1699	0·16108
3	0·1068	0·11291
4	0·0781	0·07751
5	0·0552	0·05253
6	0·0384	0·03530
7	0·0255	0·02358
8	0·0193	0·01567
9	0·0098	0·01038
10	0·0015	0·00686
11	0·0141	0·00452
12	− 0·0006	0·00297
13	− 0·0065	0·00195
14	0·0109	0·00128
15	− 0·0004	0·00084
16	− 0·0038	0·00055
17	0·0035	0·00036
18	− 0·0005	0·00023
19	− 0·0006	0·00015
20	0·0004	0·00010
21	0	0·00006
≥ 22	0	0·00012
Σ	1·0000	1·00000

CHAPTER 5.

Survey of the Theory of Cosmic Radiation.

§ 5.1. Cosmic radiation has for rather a long time been known experimentally to consist of two clearly distinguishable parts called, respectively, the *soft* and the *hard component*[1]). The reason for this terminology is that the former is completely absorbed in 10 cm of lead whereas the latter can penetrate even several metres of lead. The soft component consists of very energetic *electrons*[2]) and light quanta, called *photons*. The hard component consists essentially of a new kind of particles hitherto not found in laboratory experiments. They are of intermediate mass, about 200 times as heavy as electrons, and like these carry positive and negative charges of the same magnitude as electrons and protons. They are now called *mesons*. Theoretically, *neutral* mesons ought also exist, but such particles do not seem to have been observed[3]). Besides the mesons the hard component also contains, although probably to a small degree only, protons and neutrons and perhaps also atomic nuclei of heavier elements.

The theory of the mesons has been the centre of theoretical interest in recent years, but although much progress has been made[4]), the theory has not yet attained such a form as to allow of a quantitative theory of the secondary effects of mesons in cosmic radiation. The

---

[1]) For a survey of our present knowledge, theoretically as well as experimentally, of cosmic radiation we may e. g. refer the reader to ARLEY (1940), BRADDICK (1939), EULER und HEISENBERG (1938), FROMAN and STEARNS (1938) and HEITLER (1939).

[2]) By this term we understand, as already mentioned in chap. 4, both *positive* and *negative* electrons. The former we call *positrons* and the latter *negatrons*.

[3]) The conclusion previously drawn by ARLEY and HEITLER (1938) as to the existence of neutral mesons now appears to have been based on a less reliable experiment. Cf. also the negative results of LOVELL (1939) and TRUMPY and BJORDAL (1942).

[4]) Among the many papers on this subject we shall quote only MØLLER and ROSENFELD (1940) in which the literature in question is also quoted.

effects of the *hard* component have, therefore, to be subtracted experimentally when experiments are compared with theory. On the other hand, the theory of the *soft* component has now developed to such a degree as to give not only a qualitative but in several cases even a quantitative correct explanation of the experimental material. *We shall in this chapter give a short survey of this theory of the soft component.*

The most characteristic phenomenon of the soft component consists in the formation of *showers*, which we have already discussed in § 4.1. This phenomenon it was for a long time quite impossible to understand theoretically because it was thought that all the particles in a shower were created in one single, explosion-like, elementary process. Such a process was, however, so unlikely to occur that it would never be observed. It is now known from Wilson-chamber photographs that the showers, as a rule, are not of this 'explosion' type but of the *cascade* or *multiplication* type we have described in § 4.1. In fig. 9 we show three of the many beautiful

a b c
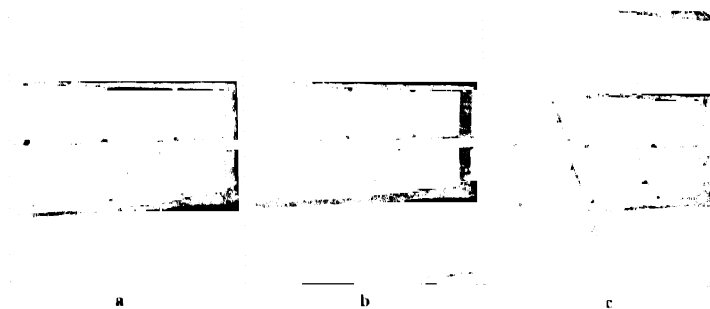
Fig. 9a, b, c. *Three typical shower photographs taken by Fussell and Street showing the growth of the showers in three divided lead plates placed across the Wilson chamber. It will be noted that in* (a) *the energy is transferred from the first to the second plate by a non-ionizing ray.*

photographs (taken by Fussell and Street) showing how the shower is multiplied up in each of the three thin lead plates placed across the chamber.

The theory of the showers was first given simultaneously by Bhabha and Heitler[1]), and Carlson and Oppenheimer[2]) and has

[1]) Bhabha and Heitler (1937). We shall in the following quote this fundamental paper as B & H.

[2]) Carlson and Oppenheimer (1937).

later on formed the subject of several other papers[1]). The theory in the form of B & H, which to the author's mind is to be preferred, has been enlarged upon by the author in two papers[2]) especially dealing with the theory of those experiments on cosmic rays using the *method of coincidence between* GEIGER-MÜLLER *counters*. The classical form of these experiments is shown in fig. 10, in which *s* denotes the shower producing layer of e. g. lead, and *a*, *b* and *c* are
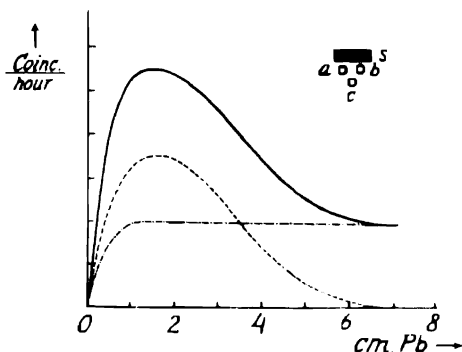
FIG. 10. *Illustration of the classical experiment of Rossi. The full curve gives the total intensity, the dotted curve that part due to the soft and the stippled one that part due to the hard component.*

three GEIGER-MÜLLER tube counters in coincidence coupling. The number of coincidences pr. unit of time, e. g. hour, is then measured and plotted against varying thicknesses of the layer. The curve thus obtained, a socalled **Rossi-curve**, is shown by the full curve in fig. 10. Now, it will be seen that in this classical form the resulting effect of both electrons, photons and mesons is measured. That part of the curve which is due to the soft component, electrons + photons, is indicated by the dotted curve, the 'soft' curve, whereas the stippled curve indicates that part, the 'hard' curve, which is due to the hard component, the mesons. We shall in the next chapter return to the question of how this experiment has to be refined in order to isolate the effects of the two components (cf. § 6.5).

[1]) BRUINS (1941). EULER (1938). FURRY (1937). HEITLER (1937). IWANENKO and SOKOLOW (1938). LANDAU and RUMER (1938). NORDHEIM (1938). NORDSIECK, LAMB and UHLENBECK (1940). SERBER (1938). SNYDER (1938).

[2]) ARLEY (1938) quoted in the following as A.
ARLEY and ERIKSEN (1940) quoted in the following as A & E.

We shall now outline how the 'soft' part of the Rossi-curves may be calculated theoretically. First, by means of the B & H theory one calculates, as we shall show in the next paragraphs, the mean number $\overline{N}(x, E_0)$, or $\overline{N}(x, k_0)$, of secondary electrons which are produced in the multiplication shower by a primary electron, or photon, of energy E_0, or k_0, having passed through a certain thickness of the material, x. It is, namely, known both theoretically and from Wilson-chamber photographs that in the first approximation the direction of motion is conserved in the shower (cf. fig. 9) and *we may, therefore, treat the problem as a one-dimensional problem.* Next one calculates the probability, $P_{el}(N, \overline{N}(x, E_0))$ and $P_{ph}(N, \overline{N}(x, k_0))$ respectively, of finding exactly N electrons in a shower, produced by an electron or a photon respectively, in which the mean value $\overline{N}$ is given. For from fig. 10 we see that in the experimental arrangement in question only showers containing *at least two* electrons can be counted (this number may, of course, be different in other experimental arrangements). *It is just this* **fluctuation problem** *of finding the probabilities of N for given $\overline{N}$ which is the subject of the present investigation.*

Finally, we have to take into account the fact that during the measurements, showers are counted which have been produced by electrons and photons of widely different energies. What is measured in coincidence experiments is, therefore, the *mean* effect produced by all the incident radiation impinging on the material during the time of investigation. We have, consequently, finally to calculate the *mean probabilities*, $\overline{P}(N, x)$, with respect to the energy spectrum. Let us by $F_{el}(E_0)dE_0$ and $F_{ph}(k_0)dk_0$ denote the probabilities of an electron and a photon having energies between E_0 and E_0+dE_0 and k_0 and k_0+dk_0, respectively, at the place of the experiment. Let us, furthermore, by p_{el} and p_{ph} denote the probabilities of a 'soft' particle being an electron and a photon, respectively. These factors will, of course, partly depend on the nature of the soft component and partly on the experimental arrangement. We thus have to calculate

$$\overline{P}(N, x) =$$
$$p_{el}\int P_{el}(N, \overline{N}(x, E_0))F(E_0)dE_0+p_{ph}\int P_{ph}(N, \overline{N}(x, k_0))F(k_0)dk_0 \quad (1)$$

giving the mean probabilities of finding N electrons in a shower emerging from a layer of thickness x.

Now the Rossi-curves measured will, obviously, be determined, apart from a proportionality factor representing the intensity of the incoming radiation, the dimensions of the layer and so on, firstly by the mean probabilities of N electrons emerging, as given by (1), and secondly by the probabilities, $g(N)$, the *geometrical factors*, of these N particles producing a coincidence:

$$\text{number of coincidences pr. hour} = \text{constant} \cdot \sum_N g(N)\bar{P}(N, x) . \quad (2)$$

These geometrical factors depend on the angular spread in the showers, the geometrical arrangement of material and counters, the sensitivity of the counters, and so on. As discussed in A[1]) these factors may, presumably, although strictly speaking they are unknown, be put equal to one except for very special experimental arrangements (cf. however § 6.5).

The nature of the 'hard' part of the Rossi-curves is still an open question. Although it is beyond the scope of the present paper to discuss this problem in detail we shall briefly mention it. The 'hard' part of the Rossi-curves now seems to be due to the following process. By their passage through matter the mesons ionize and by this process they may give rise to energetic secondary electrons, which again give rise to ordinary multiplication showers, socalled *knock-on showers*. For a time it was believed that the mesons might also give rise to photons which again produced multiplication showers and, furthermore, that the mesons might produce showers consisting of mesons. Now the probabilities of both these processes seem to be extremely small, if they occur at all[2]). Moreover, the much disputed second maximum of the Rossi-curves at about 17 cm Pb, which was interpreted in terms of meson-showers, does not seem to exist[3]). This latter fact is in agreement with the above interpretation of the 'hard' Rossi-curves in terms of knock-on showers as the resulting curves will just show *saturation form* because the primary radiation, the mesons, is far more penetrating than the secondary one, the electrons and photons. BHABHA[4]) has calculated the probabilities of these knock-on showers, and his theory seems to be in agreement with the above mentioned experiments of LOVELL[2]). BHABHA's theory applies, however, only to thicknesses so large that the mesons are in equilibrium with their secondary electrons and, furthermore, it is based on a form of the meson theory which is now abandoned[5]).

[1]) Cf. A p. 534.

[2]) Cf. especially the investigations of LOVELL (1939).

[3]) ROSSI and JÁNOSSY (1939 b). NERESON (1942).

[4]) BHABHA (1938).

[5]) The author intends to investigate these knock-on showers for smaller thicknesses on the basis of the recent form of the meson theory.

§ 5.2. We shall now give B & H's deduction of the mean number, $\overline{N}_{el}(x, E_0)$ and $\overline{N}_{ph}(x, k_0)$, of electrons produced by a primary electron and photon of energy E_0 and k_0, respectively, in a layer of thickness x[1]). The quantum theory of radiation[2]), firstly, shows that the theory is, in the first approximation, *symmetric* in positrons and negatrons. Secondly, it shows that for relativistic energies, i. e. 10^6 e. v. and upwards, only three elementary processes are of importance for the soft component, namely '*bremsstrahlung*' and *ionization* for the electrons and *pair production* for the photons. All other processes such as the Compton and the photoelectric effects may be neglected in the first approximation. Furthermore, 'bremsstrahlung' and ionization may, in the first approximation, be treated separately. This is the case because the former is the only important one for very high energies while it may be considered negligible compared with the other as regards small, although still relativistic, energies. This fact is taken into account by introducing for each material a certain **critical energy**, denoted by E_c, defined as that energy at which an electron loses just as much energy by bremsstrahlung as by ionization. This energy is approximately given by

$$E_c \sim \frac{1600mc^2}{Z} = \frac{8 \times 10^8}{Z} \text{ e. v.} \tag{1}$$

Here $mc^2 = 0{\cdot}51 \times 10^6$ e. v. is the rest energy of an electron and Z the atomic number of the material. In table 11 we give the values of E_c for various materials.

Table 11. *The critical energy, E_c, for various materials.*

	Pb	Fe	Al	H_2O	atm. air at *n. p. & t.*
E_c in e. v.	1×10^7	3×10^7	6×10^7	15×10^7	15×10^7

***The approximation now consists in the assumption that* above E_c *an electron can lose energy only by 'bremsstrahlung' and* below E_c *only by ionization*.** We first outline the calculation of the mean number of electrons with energies *above* E_c. From the quantum

[1]) Cf. B & H, A and A & E.

[2]) Cf. e. g. the textbook by Heitler (1936).

theory of radiation it is known, firstly, that *the probability of an electron with energy E_0 emitting a photon with energy between k and $k+dk$ in travelling a distance dx* is approximately given by

$$P_{ph}dkdx = c\sigma Z^2 \frac{\ln 2\, dk\, dx}{k} \quad (c\sigma Z^2 \text{ has the dimension cm}^{-1}) \qquad (2)$$

in which c is a constant, σ the number of nuclei pr. cm^3 and Z the atomic number of the material. Secondly, *the probability of a photon with energy k_0 being absorbed while travelling a distance dx thereby emitting an electron pair in which the positron has an energy between E_+ and E_++dE_+, the negatron thus lying between k_0-E_+ and $k_0-E_+-dE_+$*, is approximately given by

$$P_{pair}dE_+dx = ac\sigma Z^2 \frac{dE_+}{k_0} dx \qquad (3)$$

in which a is a dimensionless constant $\sim 0{\cdot}6$ and c, σ and Z have the same meaning as in (2). Thirdly, *the probability of an electron with energy E_0 retaining an energy greater than E after having traversed a distance x* is approximately given by

$$W\left(c\sigma Z^2 x, \ln\frac{E_0}{E}\right) = W(l, y) = \int_0^y \frac{t^{l-1}e^{-t}}{(l-1)!} dt. \qquad (4)$$

Here W is the incomplete Γ-function[1]) and

$$l = c\sigma Z^2 x\,, \qquad (5)$$

$$y = \ln\frac{E_0}{E} \qquad (6)$$

are two *dimensionless* variables. We now see that if we measure all lengths in the dimensionless *shower units* l given in (5), *the theory becomes independent of the material.* In table 12 we give the lengths x_0 corresponding to $l = 1$ for various materials.

Table 12. *The shower-unit of length, x_0, for various materials.*

	Pt	Au	Pb	Sn	Cu	Fe	Al	C	H_2O	atm. air at *n. p. & t.*
x_0 in cm	0·207	0·215	0·358	0·824	1·02	1·26	6·71	15·5	34	27500

[1]) Cf. the table of PEARSON (1922).

This shower unit x_0 has, by the way, an immediate physical interpretation as it gives *the distance in which an electron loses on the average half its energy due to 'bremsstrahlung'*. From (4) and (6) we find, in fact, that the mean energy loss in the distance $l = 1$ is given by

$$\overline{E_0 - E} = \int_0^{E_0} E \cdot dW = \int_0^{\infty} E_0 e^{-y} \cdot e^{-y} dy = \frac{E_0}{2}. \tag{7}$$

Introducing the shower unit (5) into (2) and (3) we have

$$P_{ph} dk dl = \ln 2 \frac{dk}{k} dl \tag{8}$$

and

$$P_{pair} dE_+ dl = \alpha \frac{dE_+}{k_0} dl. \tag{9}$$

Integrating (8) we find for *the total probability of an electron emitting a photon* (with energy above E_c[1]))

$$Q_{ph} dl = \ln 2 \, dl \int_{E_c}^{E_0} \frac{dk}{k} = \ln 2 \ln \frac{E_0}{E_c} dl = 0{\cdot}7 \; y_c dl. \tag{10}$$

Integrating (9) we find, due to our assumption $k_0 \gg mc^2$, for *the total probability of a photon emitting an electron pair*

$$Q_{pair} dl = \alpha dl \int_{2mc^2}^{k_0} \frac{dE_+}{k_0} = \alpha dl = 0{\cdot}6 \; dl. \tag{11}$$

(10) *and* (11) *show that the assumptions underlying the simplified stochastic model of the multiplication process, which we have discussed in* §§ 4.7–4.8 *and which we shall apply in the next chapter, are fairly correct, the probabilities mentioned being in fact of the same order of magnitude and both proportional to dl. There is, however, an important difference because here* (10) *gives the probability of just one photon being emitted, the electron still living afterwards, whereas in our model it gives the probability of two photons being emitted, the electron thereby dying. This assumption was, on the one hand, necessary in the model in order to make it symmetric in electrons and photons which fact was again necessary in order to make the model so simple that the moments might be evaluated. On the other hand, this assumption is also physically quite plausible, firstly, because in reality an electron loses its energy in*

---

[1]) We note that we cannot integrate to 0 due to the divergence of $\int_0 \frac{dk}{k}$. This is the socalled 'ultra-red catastrophe'.

the form of a few photons rather than in emitting one single photon and when the energy of the electron has decreased very much we can simply consider it as being 'dead' because it will then neither enter into the calculations nor be observed experimentally. Secondly, the theory shows that the number of electrons and photons in a shower is in fact practically identical. Finally, as regards our third assumption of both electrons and photons having a certain probability of 'dying', this assumption had to be introduced in order to account for the energy degradation taking place in the shower (cf. p. 110). A particle in the model 'dying' thus simply corresponds to the particle in the shower having so little energy that it will, as just mentioned, neither enter into calculations nor into observations. On the whole, we thus conclude that we think our simplified model in § 4.8 a fairly close approximation, reproducing, in any case, the main features of the multiplication process.

We can now start the calculations. From (11), which gives the absorption coefficient of the photons, we find that the probability, p, of a photon traversing a thickness l is given by

$$-dp = p(l)-p(l+dl) = p(l) \cdot a dl \tag{12}$$

whence

$$p(l) = e^{-al} . \tag{13}$$

We now denote by

$h(l, k)dk$ *the average number of photons which at a depth l below the surface have energies between k and $k+dk$.*

$f_{\pm}(l, E)$ *the average number of positrons and negatrons, respectively, which at a depth l below the surface have energies above E.*

We first assume that the shower is initiated by one primary positron with energy E_0. Using (8) and (13), $h(l, k)dk$ is then given by

$$h(l, k)dk = \int_0^l \{f_+(l', k)+f_-(l', k)\} \cdot \frac{\ln 2\, dk\, dl'}{k} \cdot e^{-a(l-l')} . \tag{14}$$

Using (9) and (4), $f_-(l, E)$ is given by

$$f_-(l, E) = \int_0^l \int_E^{E_0} \int_{E'}^{E_0} h(l', k)dk \cdot a \frac{dE'}{k} dl' \cdot W\left(l-l', \ln \frac{E'}{E}\right) \tag{15}$$

and $f_+(l, E)$ by

$$f_+(l, E) = f_-(l, E) + W\left(l, \ln \frac{E_0}{E}\right) . \tag{16}$$

Inserting (14) and (16) into (15) we obtain finally the *continuity equation* for $f_{-}(l, E)$

$$f_{-}(l, E) = \int_0^l dl' \int_E^{E_0} dE' \int_{E'}^{E_0} dk \int_0^{l'} dl'' \cdot$$

$$\left\{2f_{-}(l'', k) + W\left(l'', \ln\frac{E_0}{k}\right)\right\} \frac{\ln 2}{k}\, e^{-\alpha(l'-l'')} \frac{\alpha}{k} W\left(l-l', \ln\frac{E'}{E}\right). \tag{17}$$

This is the equation we have to solve. We shall not here repeat the beautiful calculations of B & H who showed that this equation can in fact be solved and that exactly! Introducing the logarithmic variable y given in (6) and writing simply $f(l, y)$ for $f(l, E_0 e^{-y})$ B & H show that the exact solution of (17) is given by

$$f_{-}(l, y) = f_{+}(l, y) - W(l, y) = \sum_{n=1}^{\infty} f_n^{(el)}(l, y) \tag{18}$$

$$f_n^{(el)}(l, y) = \left(\frac{2\alpha \ln 2}{2}\right)^n e^{-\alpha l} \int_0^l dl' \int_0^y dy' e^{\alpha l'} \frac{l'^n (l-l')^{n-1} (y-y')^{n-1}}{n!\,(n-1)!\,(n-1)!} W(l'+n, y')$$

(positron-initiated shower).

For a negatron-initiated shower we find, of course, the same expression with f_{-} and f_{+} interchanged in (18). By an analogous calculation A & E find that for a shower initiated by one primary photon with energy k_0 the exact solution of the corresponding continuity equation is given by

$$f_{-}(l, y) = f_{+}(l, y) = \sum_{n=1}^{\infty} f_n^{(ph)}(l, y)$$

$$f_1^{(ph)}(l, y) = \alpha e^{-\alpha l} \int_0^l dl' e^{\alpha l'} W(l'+1, y) \tag{19}$$

$$f_n^{(ph)}(l, y) = \frac{(2\alpha \ln 2)^n}{2 \ln 2} e^{-\alpha l} \int_0^l dl' \int_0^y dy' e^{\alpha l'} \frac{l'^{n-1} (l-l')^{n-1} (y-y')^{n-2}}{(n-1)!\,(n-1)!\,(n-2)!} W(l'+n, y')$$

(photon-initiated shower).

For proof of (18) and (19) we refer to B & H and A & E. We shall only here point out that the individual terms in the series, $f_n(l, y)$, have an immediate physical interpretation as *the average number of secondary electrons in the n'th 'generation'*, i. e. those electrons which have been produced by $2n$ transformations electron → photon and photon → electron from the 'parent electron', in the case of electron-initiated showers, and by $2n-1$ such transformations from the

'parent photon', in the case of photon-initiated showers. This fact is, of course, utilized for the proof which, therefore, runs by iteration.

The functions (18) and (19) have now been tabulated for all values of l and y which may be met with. The results are given in B & H, A and A & E. For the total average number of secondary electrons with energies above E_c, which we have called **fast** electrons and denoted by $\overline{N}_f$, we finally have

$$\overline{N}_f = 2f(l, y_c) \tag{20}$$

with

$$y_c = \ln \frac{E_0}{E_c} \quad \text{and} \quad y_c = \ln \frac{k_0}{E_c} \tag{21}$$

in case of electron- and of photon-initiated showers, respectively.

§ 5.3. In the preceding paragraph we have outlined the calculation of the *fast* secondary electrons, i. e. electrons having energies *above* the critical energy E_c. In many counter experiments the measurements include, however, also the shower electrons having energies *below* E_c. We therefore have to estimate the average number of these **slow** electrons, as we have called them. They may, under the general assumptions mentioned in the preceding paragraph, be produced in two different ways. First, they can be produced *by the slowing down of the fast electrons due to ionization* (I). Next, they can be produced *by a photon emitting an electron pair one of which has an energy below* E_c (II). We shall now shortly outline the calculations of these two kinds of secondary electrons which we have denoted by $\overline{N}_s$(el) and $\overline{N}_s$(phot), respectively (cf. A and A & E).

(I) We first discuss $\overline{N}_s$(el) (cf. A). In order that an electron should emerge with an energy below E_c it must once have had an energy greater than E_c and lost a photon with so high an energy that it retains an energy less than E_c. The probability of an electron with energy $E'(> E_c)$ emitting a photon with energy between k and $k+dk$ in travelling a distance dl is given by (5.2.8). The probability of an electron of energy E losing an energy between E'' and $E''+dE''$ due to ionization in travelling the distance l'' we denote by $P_{ion}(E'', l'')dE''$. Now, the mean energy loss due to ionization may be regarded as a constant, β, even down to energies of the order of 100,000 e. v.

$$-\left(\frac{dE}{dl}\right)_{ion} = \beta = \ln 2 \cdot E_c \tag{1}$$

which equation follows from (5.2.8) and the definition of E_c. We thus have,

using the DIRAC δ-function and the Δ-symbol introduced by the author, (1.2.10),

$$P_{ion}(E'', l'')dE'' = \Delta(E \leq E_c)\Delta(E'' \leq E)\delta(E''-\beta l'')dE'' . \tag{2}$$

We now denote by $p(l', E')dE'$ the average number of electrons which at l' have energies between $E'(> E_c)$ and $E'+dE'$ which function is simply given by differentiating the total electron spectrum $2f(l, E)$. We then have the average number of slow electrons produced by the slowing down of the fast electrons due to ionization, $\overline{N}_s$(sec. el), given by

$$\overline{N}_s(\text{sec. el}) = \int_0^l dl' \int_{m\epsilon^2}^{E_c} dE \int_{E_c}^{E_0} dE' \int_0^{E'} dk\, p(l', E')P_{ph}(E', k)P_{ion}(E'-k-E, l-l') . \tag{3}$$

Using (5.2.8), (1) and (2) A and A & E find that this expression may be written

$$\overline{N}_s(\text{sec. el}) = 2\ln 2 \int_\gamma^l \left(\frac{\partial f(L, y)}{\partial y}\right)_{y=y_c} dL$$

with $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (4)

$$\gamma = \max\left\{l - \frac{1}{\ln 2} = l - 1{\cdot}44,\ 0\right\}$$

for both electron- and photon- initiated showers. In the case of the former the primary electron may contribute in the same way producing an average number $\overline{N}_s$(prim. el) given by

$$\overline{N}_s(\text{prim. el}) = \ln 2 \int_\gamma^l \frac{y_c^{L-1} e^{-y_c}}{(L-1)!} dL \tag{5}$$

with γ given in (4).

(II) We now finally discuss $\overline{N}_s$(phot) (cf. A & E). The slow electrons of this second type may be produced either by the primary photon, in photon-initiated showers, or by the secondary photons, in both electron- and photon-initiated showers. The former slow electrons we have denoted by $\overline{N}_s$(prim. phot) and the latter by $\overline{N}_s$(sec. phot). The probability of the primary photon with energy k_0 penetrating to a depth l below the surface is given by (5.2.13). *This formula is independent of the critical energy* E_c, as this energy plays no rôle in the pair production, which process is predominant down to much lower energies, of the order 10^6 e. v. in lead. The expression (5.2.9) for the probability of pair production is, however, only valid for larger values of k_0. For smaller energies it is reduced by a factor $\theta(k_0)$, which depends on the energy k_0 (being, of course, 0 for $k_0 \leq 2mc^2$) and the shower producing material

$$P_{pair}dE_+dl = \sigma\theta(k_0)\frac{dE_+}{k_0}dl \tag{6}$$

$$(0 \leq \theta \leq 1, \quad \theta = 0 \text{ for } k_0 \leq 2mc^2, \quad \theta = 1 \text{ for } k_0 \gg 2mc^2) .$$

In spite of this fact (5.2.13) roughly holds good with a *constant* value of the absorption coefficient σ through the *whole* energy region. If, namely, the Compton and the photoelectric effects are also taken into consideration, it is found that the *total* absorption coefficient remains roughly constant through

the whole energy range considered and is equal to the value of the absorption coefficient due to pair production for high energies, which is just a. Using, furthermore, (2) we thus have for the average number of *positrons created* in this way

$$\bar{N}_s^+ \text{ (prim. phot)} =$$

$$\int_0^l dl' \int_{mc^2}^{\min} dE \int_{mc^2}^{k_0} dE_+ e^{-al'} \cdot \frac{a}{k_0} \theta(k_0) \cdot \delta(E_+ - E - \beta(l-l')) \Delta(E_+ \leqq E_c) \tag{7}$$

$$\min = \min \{k_0, E_c\}.$$

This expression may be evaluated elementarily but the result, given and tabulated in A & E, is complicated and we shall not quote it here. If the energy k_0 of the primary photon is greater than $2E_c$ and if one of the twin electrons is produced with an energy *below* E_c, then the other electron is bound to have an energy *above* E_c. In this case (7) gives the total average number of electrons created in this way. In case $k_0 < E_c$, both twin electrons will have energies *below* E_c and the total average number of electrons is, therefore, given by (7) multiplied by a factor 2. In the range $E_c < k_0 < 2E_c$ *both* electrons may happen to have energies *below* E_c or one may have an energy *below*, the other *above* E_c, but both cannot have energies above E_c. These facts are taken into account roughly in the following way. For $k_0 > E_c$ we take $\theta(k_0) \equiv 1$ and the total average number of electrons equal to $\bar{N}_s^+$(prim. phot) as given in (7). For $k_0 < E_c$ we take the total average number of electrons equal to $2\bar{N}_s^+$(prim. phot) and for $\theta(k_0)$ we take the value given by the quantum theory of radiation[1]).

In the same way we can proceed for the secondary photons. Using the expression for $h(l, k)dk$ given in (5.2.14)[2]) we have, corresponding to (7),

$$\bar{N}_s^+ \text{ (sec. phot)} = \tag{8}$$

$$\int_0^l dl' \int_{mc^2}^{E_c} dE \int_{2mc^2}^{k_0} dk \int_{mc^2}^{k} dE_+ h(l', k) \cdot \frac{a}{k} \theta(k) \cdot \delta(E_+ - E - \beta(l-l')) \Delta(E_+ \leqq E_c).$$

This expression may be somewhat reduced. The result, given and tabulated in A & E, is again rather complicated and we shall not quote it here. We shall only note that the expression (8) may be split into two parts representing the contribution from such secondary photons as have energies, respectively, above and below the critical energy E_c. In order to obtain the total average number of electrons produced in this way we proceed as in the case of the primary photon just discussed above. In the first term we put $\theta \equiv 1$ and do not multiply by 2 whereas in the second term we take the value for $\theta(k)$ as discussed above and multiply by 2. By this procedure the first and second term becomes, respectively, independent and dependent of the shower producing material and we therefore write

$$\bar{N}_s\text{(sec. phot)} = \bar{N}_s\text{(sec. phot. ind)} + \bar{N}_s\text{(sec. phot. dep)}. \tag{9}$$

[1]) Cf. e. g. fig. 18 p. 201 in HEITLER (1936).

[2]) We note that for $k < E_c$ we shall as a consequence of our assumptions substitute $f_+(l'', E_c) + f_-(l'', E_c)$ for $f_+(l'', k) + f_-(l'', k)$ in (5.2.14).

We have now outlined the calculation of the average numbers of both 'fast' ($E > E_c$) and 'slow' ($E < E_c$) secondary electrons. All the functions have been calculated for all values of the two shower variables l and y_c and the results are tabulated in B & H, A and A & E. In tables 13 (p. 152) and 15 (p. 159) are given the total average numbers of fast+slow secondary electrons as functions of l and $y_c = \ln \frac{k_0}{E_c}$ and $y_c = \ln \frac{E_0}{E_c}$ for photon- and electron-initiated showers, respectively. In fig. 4 p. 110 we have shown the former curves as functions of l for different values of y_c.

Finally, we observe that the total number of 'slow' electrons calculated as outlined in the present paragraph will presumably exceed possible experimental values somewhat since these electrons will have rather a great chance of being scattered in the material. As a consequence we ought, therefore, perhaps include a certain fraction only of the 'slow' electrons in the subsequent calculations. However, on the one hand, it is extremely difficult to treat this scattering in a proper way and, on the other hand, it has turned out that the *form* of the resulting Rossi curves (5.1.2) does not depend significantly on the number of slow electrons included.

CHAPTER 6.

Application of Stochastic Processes to the Theory of Cosmic Radiation. Discussion of Results and Comparison with Experiments.

§ 6.1. As we have outlined in the preceding chapter the theory of showers allows of the direct deduction of *continuity equations* for the mean numbers of electrons (and of photons). *This is, however, impossible for the corresponding fluctuations.* In order to deduce the continuity equations for the second moments we must have some knowledge of the probability distribution itself. Now we have seen that the calculation of the mean values is already so complicated that it would be quite hopeless to try to evaluate directly the real probability distribution and thus the second moments. We therefore now have to utilize our stochastic model discussed in § 4.8. *The guiding principle is here to evaluate for each value of l and y_c what is the fluctuation of the model when this is properly adjusted to give the closest possible picture of the shower in question.* With the fluctuation thus calculated we can at once, using the Pólya approximation as discussed in § 4.9, obtain reliable estimates of the probabilities $P_{el}(N, l, E_0)$ and $P_{ph}(N, l, k_0)$, whose mean values with respect to the energy spectra of E_0 and k_0 just give the Rossi curves we are looking for, as discussed in § 5.1.

We first discuss the case of *photon-initiated showers* which is the simplest case as all the electrons emerging are *secondary* ones. The adjustment of the model to a given shower, i. e. with given l, y_c and material[1]), may, of course, be done in several ways. As we remember, the variable y_c $\left(= \ln \frac{k_0}{E_c}\right)$ is a measure of the energy of the primary photon and thus also of the penetrating power of the secondary

[1]) For the sake of simplicity we shall only discuss showers from *lead* as the fluctuation problem depends only to a lesser degree on the material.

photons and electrons. On the other hand, the parameter γ of the model is a measure of the 'death rate' of the particles in the model, i. e. of their penetrating power. It is, therefore, a natural idea to adjust the model in the following way. For each fixed value of y_c we adjust by γ means of (4.8.10)[1]) to give the same maximum value of the mean number $\bar{m}$, given in (4.8.8)[2]), in the model as of the mean number $\bar{N}$ of electrons in the shower. y_c and γ being now fixed we next compare the two curves $\bar{N}$ and $\bar{m}$ as functions of l and t, respectively. For each value of l we determine from the curves the value of t giving the same mean number and lying on the same side of the maximum as the l point

$$\bar{N}(l) = \bar{m}(t) \ . \tag{1}$$

If the model were a very close approximation, the corresponding l and t values would be proportional, the factor of proportionality being just the second parameter λ measuring the 'birth rate' of the model. In fact, we do not find strict proportionality as shown in fig. 11 giving $\lambda = \frac{l}{t}$ (for lead) as a function of l for different values of y_c. It is, however, seen from the fig. that λ is fairly constant for
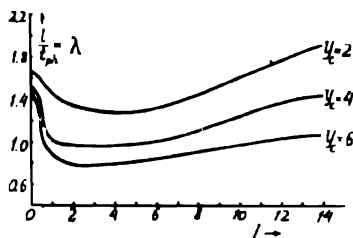
FIG. 11. *The ratio of the shower thickness l and the thickness t of the adjusted model for photon-initiated showers in lead.*

each value of y_c which fact shows that the approximation is reasonably good[3]). The value of both γ and t being now determined, we

[1]) The other parameter λ, which is a measure of the 'birth rate', we put equal to 1 for the moment, i. e. we make the substitution $\lambda l \rightarrow t$ and $\frac{\gamma}{\lambda^2} \rightarrow \gamma$, cf. § 4.9.

[2]) We note that for a *photon*-initiated shower the *n*-particles of the model have to be interpreted as photons and the *m*-particles as electrons due to the initial condition (4.8.2). For electron-initiated showers it is vice versa.

[3]) It is interesting to observe that the approximating model has another feature

can calculate $\overline{m^2}$ by means of (4.8.36), $\sigma^2 = \overline{m^2} - \bar{m}^2$ and finally the second parameter b of the Pólya distribution given by (cf. (4.4.11))

$$\sigma^2 = \bar{m}(1+b\bar{m}) . \tag{2}$$

From (2) it is seen that b can be interpreted as the relative deviation of σ^2 from the square of the Poisson fluctuation, $\sigma_{Poisson} = \sqrt{\bar{m}}$. The result is given in table 13 which gives $\bar{N}$, σ^2 and b (for lead) as functions of l and y_c[1]), other values being easily obtained by graphical interpolation. In fig. 4 p. 110 we have shown the $\bar{N}$ curves for various values of y_c. In fig. 12 we show the variation of b. Remem-
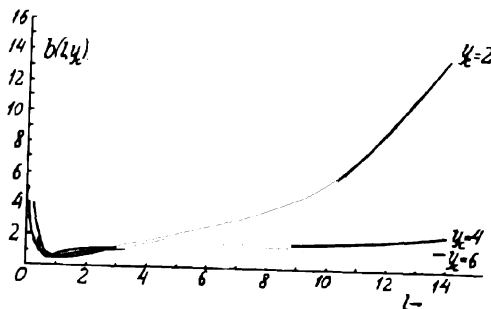
FIG. 12. *The Pólya fluctuation parameter* $b = \dfrac{\sigma^2 - \bar{N}}{\bar{N}^2}$ *(for lead; cf. table 13) giving the relative deviation of* σ^2 *from the square of the Poisson fluctuation* $\sqrt{\bar{N}}$.

bering that b measures the deviation from the Poisson distribution ($b = 0$), and gives the Furry distribution for $b = 1$, we see that neither of these distributions represent good approximations, as *the fluctuation is in fact of varying size and is in all cases* larger *than the Poisson fluctuation* $\sqrt{\bar{N}}$. We note that the fact that b increases indefinitely with increasing value of l, which fact can also be seen analytically, does not mean that the probability distribution itself

in common with the real showers. From the $\bar{N}$ curves of both photon- and electron-initiated showers (cf. tables 13 and 15) it is found that approximately $\ln \bar{N}_{max} = \alpha + \beta l_{max}$ in which l_{max} denotes the value of l giving the maximum value, $\bar{N}_{max}$, of $\bar{N}$. In fact we also have a relation of the same nature for the model, as we see from (4.8.10) that $\ln \bar{n}_{max} = \ln \bar{m}_{max} = -\ln 2 + \frac{1}{2} l_{max}$ (for $\lambda = 1$ and $l_{max} > 1$).

[1]) We note, that due to the uncertainties inherent in graphical methods we cannot expect the relation (1) and consequently, as b has been calculated from (2), nor yet $b = \dfrac{\sigma^2 - \bar{N}}{\bar{N}^2}$ to be strictly fulfilled.

Table 13. *Photon-initiated showers in lead. The mean number of electrons, $\bar{N}$, the square of the fluctuation, σ^2, and the Pólya parameter, $b = \frac{\sigma^2 - \bar{N}}{\bar{N}^2}$, giving the relative deviation of σ^2 from the square of the Poisson fluctuation, $\sigma_{Poisson} = \sqrt{\bar{N}}$, as functions of l and y_c.*

l \ y_c		2	4	6	8
0·2	$\bar{N}$	0·218	0·238	0·244	0·247
	σ^2	0·380	0·453	0·472	0·479
	b	3·36	3·95	3·78	3·67
0·4	$\bar{N}$	0·391	0·445	0·471	0·491
	σ^2	0·614	0·786	0·790	0·828
	b	1·40	1·74	1·45	1·36
0·6	$\bar{N}$	0·556	0·669	0·746	0·812
	σ^2	0·791	1·09	1·16	1·27
	b	0·692	0·580	0·731	0·485
0·8	$\bar{N}$	0·710	0·925	1·07	1·23
	σ^2	1·02	1·63	1·66	1·99
	b	0·539	0·458	0·449	0·490
1	$\bar{N}$	0·858	1·19	1·48	1·78
	σ^2	1·24	2·17	2·77	3·69
	b	0·490	0·505	0·556	0·596
2	$\bar{N}$	1·51	3·00	5·07	7·69
	σ^2	3·18	10·4	31·7	80·6
	b	0·781	0·892	1·00	1·05
3	$\bar{N}$	2·01	5·30	11·5	21·9
	σ^2	6·63	35·3	139	745
	b	1·20	1·17	1·14	1·16
5	$\bar{N}$	1·71	9·65	33·6	88·3
	σ^2	7·78	143·4	1420	9030
	b	2·07	1·44	1·30	1·19
7	$\bar{N}$	0·911	9·90	53·6	183
	σ^2	3·40	168	3820	40700
	b	3·00	1·54	1·31	1·21
10	$\bar{N}$	0·292	6·93	58·1	305
	σ^2	0·657	115	4670	114000
	b	5·45	1·73	1·35	1·22
14	$\bar{N}$	0·061	2·56	32·8	285
	σ^2	0·118	17·2	1240	99800
	b	13·9	2·27	1·42	1·23

Table 14. ***Photon-initiated showers in lead. The probability distribution $P_{ph}(N, l, y_c)$ giving the probability of a shower containing N electrons as a function of l and y_c.***

N \ y_c		$l = 0·2$: 2	$l = 0·2$: 4	$l = 0·2$: 6	$l = 0·8$: 2	$l = 0·8$: 4	$l = 0·8$: 6
0	Pólya	0·849	0·847	0·841	0·540	0·453	0·447
	Poisson	0·804	0·787	0·783	0·484	0·386	0·330
1	Pólya	0·107	0·103	0·107	0·283	0·285	0·286
	Poisson	0·175	0·188	0·192	0·351	0·367	0·366
2	Pólya	0·0295	0·0311	0·0325	0·114	0·144	0·146
	Poisson	0·0191	0·0225	0·0235	0·127	0·175	0·203
3	Pólya	0·00956	0·0113	0·0118	0·0416	0·0665	0·0681
	Poisson	0·00139	0·00180	0·00192	0·0308	0·0553	0·0751
4	Pólya	0·00334	0·00441	0·00463	0·0143	0·0294	0·0303
	Poisson	8×10^{-5}	1×10^{-4}	1×10^{-4}	0·00557	0·0132	0·0209
≥ 5	Pólya	0·00196	0·00315	0·00330	0·00694	0·0217	0·0225
	Poisson	4×10^{-6}	5×10^{-6}	6×10^{-6}	9×10^{-4}	3×10^{-3}	6×10^{-3}

N \ y_c		$l = 1$: 2	$l = 1$: 4	$l = 1$: 6	$l = 2$: 2	$l = 2$: 4	$l = 2$: 6
0	Pólya	0·484	0·401	0·377	0·374	0·251	0·160
	Poisson	0·412	0·289	0·214	0·235	0·0539	0·00701
1	Pólya	0·295	0·280	0·266	0·257	0·194	0·138
	Poisson	0·365	0·359	0·330	0·340	0·157	0·0348
2	Pólya	0·134	0·158	0·160	0·157	0·146	0·114
	Poisson	0·162	0·223	0·254	0·247	0·230	0·0863
3	Pólya	0·0539	0·0824	0·0906	0·0920	0·108	0·0949
	Poisson	0·0478	0·0920	0·131	0·119	0·224	0·143
4	Pólya	0·0203	0·0411	0·0496	0·0528	0·0801	0·0791
	Poisson	0·0106	0·0285	0·0502	0·0432	0·163	0·177
≥ 5	Pólya	0·0135	0·0378	0·0563	0·0676	0·221	0·405
	Poisson	2×10^{-3}	0·00884	0·0204	0·0164	0·172	0·552

Table 14 cont.

	$l=3$			$l=5$		
N \ y_c	2	4	6	2	4	6
0 Pólya	0·362	0·197	0·108	0·487	0·156	0·0516
0 Poisson	0·144	0·00592	4×10^{-4}	0·196	1×10^{-4}	0
1 Pólya	0·213	0·141	0·0861	0·179	0·100	0·0397
1 Poisson	0·279	0·0304	0·00337	0·319	9×10^{-4}	0
2 Pólya	0·137	0·110	0·0742	0·103	0·0787	0·0347
2 Poisson	0·270	0·0779	0·0130	0·260	0·00411	0
3 Pólya	0·0914	0·0896	0·0655	0·0666	0·0658	0·0315
3 Poisson	0·175	0·133	0·0336	0·141	0·0127	0
4 Pólya	0·0616	0·0737	0·0584	0·0454	0·0566	0·0291
4 Poisson	0·0848	0·171	0·0651	0·0576	0·0293	0
$\geqq 5$ Pólya	0·134	0·389	0·608	0·119	0·543	0·813
$\geqq 5$ Poisson	0·0474	0·582	0·885	0·0254	0·953	1

	$l=10$			$l=14$		
N \ y_c	2	4	6	2	4	6
0 Pólya	0·853	0·215	0·0387	0·957	0·436	0·0715
0 Poisson	0·761	0·00180	0	0·944	0·0907	0
1 Pólya	0·0876	0·115	0·0284	0·0311	0·161	0·0491
1 Poisson	0·207	0·0114	0	0·0547	0·218	0
2 Pólya	0·0315	0·0840	0·0244	0·00802	0·0982	0·0409
2 Poisson	0·0301	0·0359	0	0·00159	0·261	0
3 Pólya	0·0140	0·0671	0·0220	0·00267	0·0680	0·0359
3 Poisson	0·00255	0·0757	0	3×10^{-6}	0·209	0
4 Pólya	0·00680	0·0558	0·0204	0·00099	0·0498	0·0325
4 Poisson	2×10^{-4}	0·120	0	4×10^{-7}	0·125	0
$\geqq 5$ Pólya	0·00761	0·464	0·866	0·00066	0·187	0·770
$\geqq 5$ Poisson	1×10^{-5}	0·756	1	0	0·0964	1

deviates more and more from the Poisson distribution because $b\bar{m} \to 0$ at the same time, and if $\bar{m} \ll 1$, all distributions will be very near the unit distribution $\delta_{n,0}$.

From the table of b we may finally calculate, by means of the Pólya distribution (4.4.1), the probabilities $P(N, l, y_c)$ of exactly N secondary electrons emerging from a thickness l of lead when the primary photon has the energy k_0, i. e. $y_c = \ln \frac{k_0}{E_c}$ where E_c is the critical energy of lead. In table 14 we give by way of comparison these probabilities both for the Poisson distribution previously used (cf. A & E) and for the adjusted Pólya distribution. In figs. 13–18 we show the values of table 14 graphically for a point at the beginning, middle and tail of the mean value curve. It will be seen that for small values of l the difference between the Poisson and the Pólya curves is rather small, although not quite negligible. The difference grows, however, more and more distinct with increasing values of l, especially for the higher values of y_c, i. e. the more energetic showers, the Pólya distribution giving much higher values for finding small showers (N small) than the Poisson distribution. *We should, therefore, expect to find quite different Rossi curves with the Pólya than with the Poisson distribution, especially in the tail. This last fact would be very important, as discussed in the introduction, as regards conclusions as to the existence of neutral mesons* (*cf.* § 6.5). *We shall, however, see in* § 6.3 *that this is not so because the difference is smoothed out by the averaging over the primary energy spectrum.*

§ **6.2.** We now discuss the case of *electron-initiated showers*. This case is a little more complicated than that of the photon-initiated showers because the behaviour of the primary electron in the shower is known (cf. (5.2.4)). The 'ancestor' and the 'off-spring' have, therefore, to be treated separately, the Pólya formula being applied only to the *secondary* electrons. Before we can find the t-values corresponding to the l-values we thus have to find the mean number of *secondary* n-particles, $\bar{n}_{sec}$, in the model[1]). Now the primary may disappear in two ways, either by a 'birth' or by a 'death' process, i. e. the probability, $P_{prim}(t)$, of finding the primary at t satisfies (cf. § 4.8)[2])

[1]) Cf. [2]) p. 150.

[2]) We have here again put $\lambda = 1$, i. e. performed the substitution $\lambda t \to t$ and $\frac{\gamma}{\lambda^2} \to \gamma$ (cf. § 4.9).
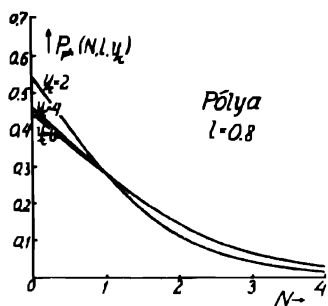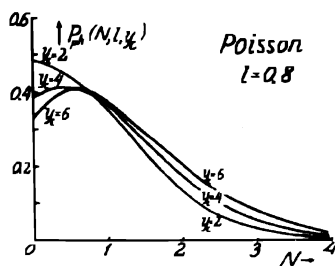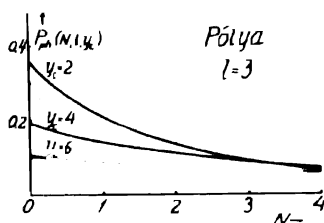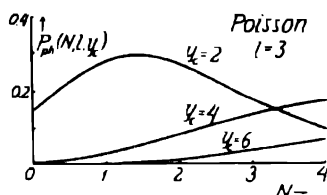
Fig. 13. Fig. 14. Fig. 15. Fig. 16. Fig. 17. Fig. 18.
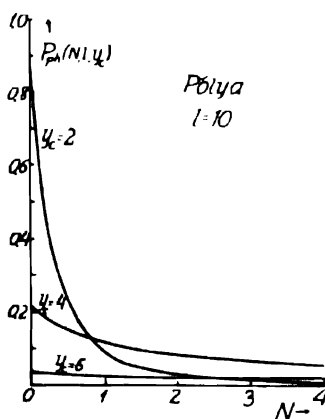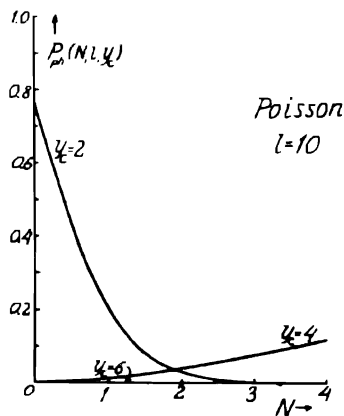
Fig. 13–18. *The Pólya and Poisson distributions $P_{ph}(N, l, y_c)$ (in lead; cf. table 14) for a point at the beginning, middle and tail of the mean value curve.*

$$\frac{d}{dt}P_{prim} = -(1+\gamma t)P_{prim} \quad (P_{prim}(0) = 1) \tag{1}$$

whence

$$P_{prim}(t) = \bar{n}_{prim}(t) = \exp\left[-(t+\frac{\gamma}{2}t^2)\right] \tag{2}$$

and[1])

$$\sigma^2_{prim} = P_{prim}(1-P_{prim}) = \exp\left[-(t+\frac{\gamma}{2}t^2)\right]\left(1-\exp\left[-(t+\frac{\gamma}{2}t^2)\right]\right). \tag{3}$$

From (4.8.8) and (2) we have

$$\bar{n}_{sec}(t) = \bar{n}(t)-\bar{n}_{prim}(t) =$$

$$\frac{1}{2}\left(\exp\left[t-\frac{\gamma}{2}t^2\right]+\exp\left[-(3t+\frac{\gamma}{2}t^2)\right]\right)-\exp\left[-\left(t+\frac{\gamma}{2}t^2\right)\right]. \tag{4}$$

If $t_{max} \gg 1$, the maximum of $\bar{n}_{sec}$ will again be given by (cf. (4.8.10))

$$(\bar{n}_{sec})_{max} = \frac{1}{2}\exp\left[\frac{1}{2\gamma}\right] \quad \text{for} \quad t_{max} = \frac{1}{\gamma}. \tag{5}$$

We now adjust our model to a given shower in the same way as for the photon-initiated showers. For each fixed value of y_c $\left(= \ln\frac{E_0}{E_c}\right)$ we adjust γ by means of (5) to give the same maximum value of the mean number $\bar{n}_{sec}$ in the model as of $\bar{N}$ of electrons in the shower. y_c and γ being now fixed, we next compare the two curves $\bar{N}$ and $\bar{n}_{sec}$ as functions of l and t, respectively. For each value of l we determine the value of t giving the same mean number and lying on the same side of the maximum as the l point

$$\bar{N}(l) = \bar{n}_{sec}(t). \tag{6}$$

As in the case of photon-initiated showers, the corresponding l and t values should be proportional. In fig. 19 we give $\lambda = \frac{l}{t}$ (for lead) as a function of l for different values of y_c and it will be seen that we have again a rough proportionality for each value of y_c[2]).

The values of both γ and t being now determined, we next have to calculate the fluctuation of the secondary n-particles, σ_{sec}, in the

[1]) Cf. e. g. A & B p. 49.

[2]) Cf. [2]) p. 150.
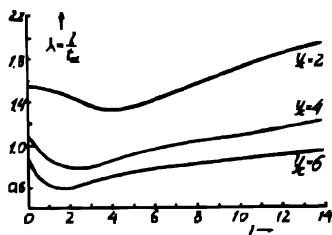
FIG. 19. *The ratio of the shower thickness l and the thickness t of the adjusted model for electron-initiated showers (in lead).*

model. This is e. g. obtained by solving the equation (4.3.29) which gives

$$\sigma_{sec} = -\varrho\sigma_{prim} + \left(\sigma^2 - (1-\varrho^2)\sigma^2_{prim}\right)^{\frac{1}{2}} \qquad (\sigma^2 = \overline{n^2} - \bar{n}^2)\,. \qquad (7)$$

Here $\bar{n}$ is given by (4.8.8), $\overline{n^2}$ by (4.8.33), σ_{prim} by (3) and ϱ is the *correlation coefficient* between the primary and the secondaries. As we have discussed in § 4.3, the distribution of the primary and the secondaries may, however, on a good approximation be regarded as being stochastically independent which fact means $\varrho \sim 0$. For our model, given in § 4.8, the corresponding calculation would be rather complicated and as the result would, of course, be the same, we shall at once put $\varrho = 0$ in (7). This procedure will be approximately correct the more so the larger t values we consider. We found, furthermore, in § 4.3 that ϱ is *negative*. The same is, of course, the case here as n_{prim} and n_{sec} vary again in opposite directions. Putting $\varrho = 0$ in (7) we may, consequently, only risk to *under*estimate σ_{sec}. Underestimating σ_{sec} we may, however, further risk that the fluctuation becomes *sub*normal, i. e. smaller than $\sqrt{\bar{n}_{sec}}$ as given by the Poisson distribution, instead of *super*normal which is the condition for the applicability of the Pólya formula. We may thus risk obtaining *negative* values of b which fact means that we cannot apply the Pólya distribution. In table 15 we give $\bar{N}_{sec}$, σ^2_{sec} and b_{sec} (for lead) as functions of l and y_0[1]), other values being easily obtained by graphical interpolation. In fig. 20 we show the variation of b_{sec}. It will be seen that we do in fact obtain negative values of b_{sec}, but only for the smallest values of l where the mean value will anyway be very small. All distributions will in this case be nearly alike and

[1]) Cf. [1]) p. 151.

Table 15. *Electron-initiated showers in lead. The mean number of secondary electrons, $\bar{N}_{sec}$, the square of the fluctuation of the secondaries, σ^2_{sec}, and the Pólya parameter, $b_{sec} = \frac{\sigma^2_{sec} - \bar{N}_{sec}}{\bar{N}^2_{sec}}$, giving the relative deviation of σ^2_{sec} from the square of the Poisson fluctuation, $\sigma_{Poisson} = \sqrt{\bar{N}_{sec}}$, as functions of l and y_0.*

l \ y_0		2	4	6
0·2	$\bar{N}_{sec}$	0·0301	0·0609	0·093
	σ^2_{sec}	0·0119	0·0321	0·0530
	b_{sec}	−19·9	−7·76	−4·61
0·4	$\bar{N}_{sec}$	0·106	0·226	0·355
	σ^2_{sec}	0·0693	0·189	0·370
	b_{sec}	−3·26	−0·690	0·098
0·6	$\bar{N}_{sec}$	0·216	0·476	0·766
	σ^2_{sec}	0·193	0·581	1·13
	b_{sec}	−0·538	0·446	0·612
0·8	$\bar{N}_{sec}$	0·343	0·789	1·31
	σ^2_{sec}	0·358	1·13	2·77
	b_{sec}	0·124	0·629	0·762
1	$\bar{N}_{sec}$	0·488	1·17	2·01
	σ^2_{sec}	0·594	2·16	5·44
	b_{sec}	0·519	0·747	0·781
2	$\bar{N}_{sec}$	1·20	3·75	7·84
	σ^2_{sec}	2·52	17·7	75·6
	b_{sec}	0·975	1·07	1·09
3	$\bar{N}_{sec}$	1·66	6·91	17·6
	σ^2_{sec}	5·32	69·0	356
	b_{sec}	1·31	1·28	1·19
5	$\bar{N}_{sec}$	1·37	10·8	45·6
	σ^2_{sec}	5·02	176	2530
	b_{sec}	2·21	1·45	1·26
7	$\bar{N}_{sec}$	0·656	10·5	70·7
	σ^2_{sec}	2·10	189	6280
	b_{sec}	3·36	1·62	1·24
10	$\bar{N}_{sec}$	0·171	5·36	65·9
	σ^2_{sec}	0·438	60	6050
	b_{sec}	9·32	1·89	1·34
14	$\bar{N}_{sec}$	0·0207	1·29	27·1
	σ^2_{sec}	0·0383	5·69	1090
	b_{sec}	35·2	2·79	1·40
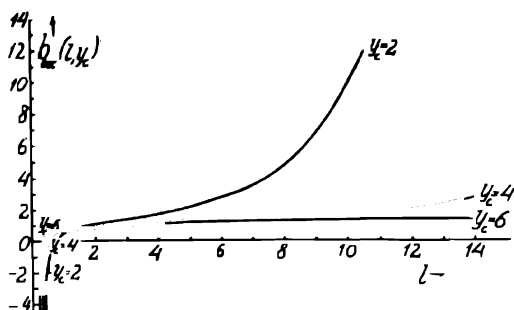
Fig. 20. *The Pólya fluctuation parameter* $b_{sec} = \dfrac{\sigma^2_{sec} - \overline{N}_{sec}}{\overline{N}^2_{sec}}$ *(for lead; cf. table 15) giving the relative deviation of* σ^2_{sec} *from the square of the Poisson fluctuation* $\sqrt{\overline{N}_{sec}}$.

we may, therefore, simply take the values of the Poisson distribution in such cases. We see again *the same trend as by the photon-initiated showers, namely, that the fluctuations are of varying size and in all cases* **larger** *than the Poisson fluctuation* ($b = 0$). Also the case $b = 1$ (the Furry distribution) would represent a bad approximation.

From the table of b_{sec} we finally calculate, by means of the Pólya distribution (4.4.1), the probabilities $P(N, l, y_c)$ of exactly N electrons, primary or secondary, emerging from a thickness l of lead when the primary electron has the energy E_0, i. e. $y_c = \ln \dfrac{E_0}{E_c}$ where E_c is the critical energy of lead. Treating the primary and the secondaries as independent (which fact involves putting $\varrho = 0$, cf. above) we have

$$\begin{aligned} P(0, l, y_c) &= \big(1 - W(l, y_c)\big)P_{sec}(0, l, y_c) \\ P(N, l, y_c) &= W(l, y_c)P_{sec}(N-1, l, y_c) + \big(1 - W(l, y_c)\big)P_{sec}(N, l, y_c) \\ & \qquad N = 1, 2, 3, \ldots \end{aligned} \tag{8}$$

in which W, given in (5.2.4), is the probability of finding the primary, and P_{sec} denotes the Pólya distribution adjusted as just discussed. In table 16 we give for purposes of comparison these probabilities (for lead) both for the Poisson distribution previously used (cf. A) and for the adjusted Pólya distribution (8). In fig. 21–26 we show the values of table 16 graphically for a point at the beginning, middle and tail of the mean value curves. *It will be seen that we here have the same trend as in the case of photon-initiated showers, the difference*
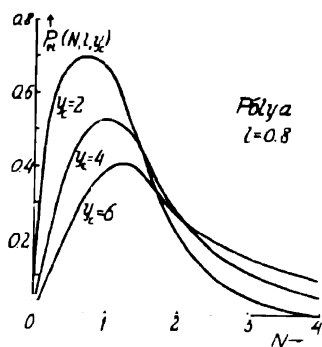
Fig. 21.
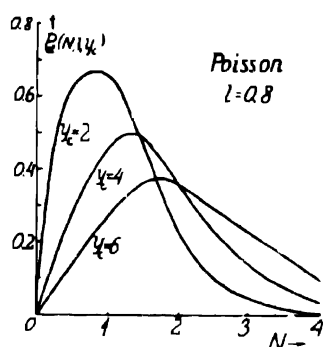
Fig. 22.
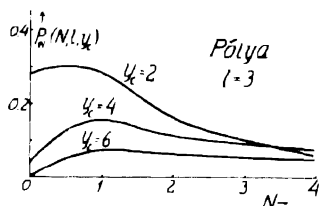
Fig. 23.
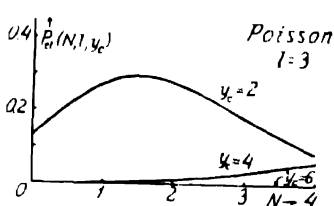
Fig. 24.
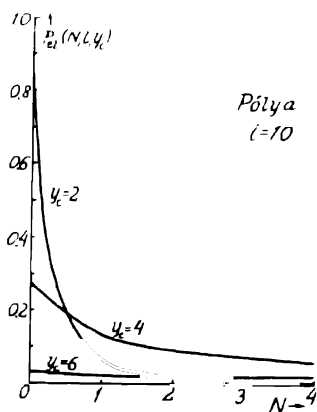
Fig. 25.
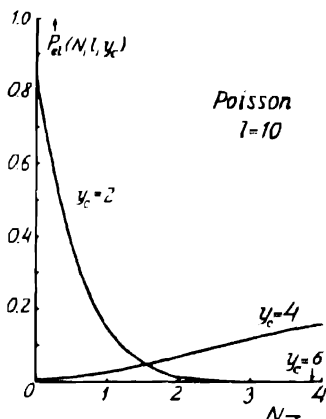
Fig. 26.

Fig. 21–26. *The Pólya and Poisson distributions $P_{el}(N, l, y_c)$ (in lead; cf. table 18) for a point at the beginning, middle and tail of the mean value curve.*

Table 16. *Electron-initiated showers in lead. The probability distribution $P_{el}(N, l, y_c)$ giving the probability of a shower containing N electrons (primary as well as secondary) as a function of l and y_c.*

		l = 0·2			l = 0·4		
N	y_c	2	4	6	2	4	6
0	Pólya Poisson	0·0126	0·0010	0·0001	0·0296	0·00255	0·00021 0·00021
1	Pólya Poisson	0·958	0·940	0·911	0·873	0·796	0·704 0·701
2	Pólya Poisson	0·0288	0·0572	0·0847	0·0924	0·180	0·243 0·249
3	Pólya Poisson	0·00043	0·00174	0·00394	0·00489	0·0203	0·0462 0·0442
4	Pólya Poisson	4×10^{-4}	$3{\cdot}5 \times 10^{-6}$	0·000122	2×10^{-4}	0·00153	0·00638 0·00523
≧ 5	Pólya Poisson	0	0	0	0	0	5×10^{-4} 5×10^{-4}

		l = 0·6			l = 0·8		
N	y_c	2	4	6	2	4	6
0	Pólya Poisson	0·0483	0·00421 0·00404	0·00043 0·00037	0·0676 0·0670	0·00616 0·00522	0·00055 0·00038
1	Pólya Poisson	0·768	0·645 0·619	0·533 0·465	0·669 0·666	0·533 0·453	0·393 0·270
2	Pólya Poisson	0·165	0·254 0·294	0·278 0·356	0·217 0·224	0·275 0·356	0·263 0·353
3	Pólya Poisson	0·0178	0·0726 0·0700	0·117 0·136	0·0400 0·0383	0·115 0·140	0·154 0·231
4	Pólya Poisson	0·00127	0·0181 0·0111	0·0453 0·0348	0·00545 0·00436	0·0448 0·0368	0·0868 0·101
≧ 5	Pólya Poisson	7×10^{-6}	0·00530 0·00146	0·0261 0·00783	0·00060 0·00040	0·0267 0·00858	0·103 0·0441

Table 16 cont.

N \ y_c		$l = 1$: 2	$l = 1$: 4	$l = 1$: 6	$l = 2$: 2	$l = 2$: 4	$l = 2$: 6
0	Pólya	0·0884	0·00793	0·00073	0·186	0·0205	0·00220
	Poisson	0·0831	0·00568	0·00033	0·122	0·00210	0·00001
1	Pólya	0·599	0·430	0·291	0·373	0·222	0·126
	Poisson	0·571	0·311	0·134	0·326	0·0215	0·00043
2	Pólya	0·226	0·267	0·230	0·204	0·165	0·103
	Poisson	0·269	0·366	0·269	0·303	0·0953	0·00325
3	Pólya	0·0651	0·145	0·163	0·110	0·127	0·0886
	Poisson	0·0648	0·210	0·270	0·164	0·166	0·0124
4	Pólya	0·0169	0·0749	0·110	0·0591	0·0991	0·0771
	Poisson	0·0105	0·0818	0·182	0·0621	0·206	0·0322
≧ 5	Pólya	0·00538	0·0746	0·206	0·0683	0·366	0·603
	Poisson	0·00141	0·0309	0·144	0·0232	0·498	0·952

N \ y_c		$l = 3$: 2	$l = 3$: 4	$l = 3$: 6	$l = 6$: 2	$l = 6$: 4	$l = 6$: 6
0	Pólya	0·279	0·0395	0·00476	0·513	0·0905	0·0116
	Poisson	0·129	0·00024	1×10^{-9}	0·241	1×10^{-5}	4×10^{-21}
1	Pólya	0·280	0·154	0·0758	0·201	0·112	0·0380
	Poisson	0·275	0·00240	5×10^{-8}	0·343	0·00015	3×10^{-20}
2	Pólya	0·158	0·111	0·0611	0·102	0·0815	0·0305
	Poisson	0·279	0·0109	6×10^{-7}	0·244	0·00083	5×10^{-18}
3	Pólya	0·0982	0·0900	0·0537	0·0615	0·0670	0·0271
	Poisson	0·183	0·0312	5×10^{-6}	0·116	0·00313	8×10^{-17}
4	Pólya	0·0628	0·0754	0·0485	0·0392	0·0574	0·0250
	Poisson	0·0876	0·0644	2×10^{-5}	0·0411	0·00886	1×10^{-16}
≧ 5	Pólya	0·121	0·530	0·756	0·0824	0·591	0·868
	Poisson	0·0466	0·891	1	0·0151	0·987	1

Table 16 cont.

N \ y_c		$l=7$: 2	$l=7$: 4	$l=7$: 6	$l=10$: 2	$l=10$: 4	$l=10$: 6
0	Pólya	0·704	0·149	0·0164	0·903	0·277	0·0315
	Poisson	0·517	2×10^{-5}	1×10^{-31}	0·843	0·00466	2×10^{-29}
1	Pólya	0·147	0·106	0·0237	0·0594	0·136	0·0262
	Poisson	0·341	0·00026	9×10^{-30}	0·144	0·0250	2×10^{-27}
2	Pólya	0·0650	0·0773	0·0201	0·0202	0·0939	0·0223
	Poisson	0·113	0·00138	3×10^{-28}	0·0123	0·0670	5×10^{-26}
3	Pólya	0·0342	0·0631	0·0183	0·00868	0·0720	0·0201
	Poisson	0·0248	0·00489	7×10^{-27}	0·00070	0·120	1×10^{-24}
4	Pólya	0·0194	0·0536	0·0171	0·00414	0·0569	0·0200
	Poisson	0·00409	0·0130	1×10^{-25}	0·00003	0·161	2×10^{-23}
≥ 5	Pólya	0·0302	0·551	0·905	0·00447	0·365	0·880
	Poisson	0·00061	0·980	1	0	0·622	1

$l=14$				
N \ y_c		2	4	6
0	Pólya	0·084	0·583	0·0727
	Poisson	0·080	0·275	2×10^{-12}
1	Pólya	0·0121	0·162	0·0505
	Poisson	0·0203	0·355	5×10^{-11}
2	Pólya	0·00270	0·0855	0·0422
	Poisson	0·00021	0·229	6×10^{-10}
3	Pólya	0·00070	0·0522	0·0372
	Poisson	0	0·0985	6×10^{-9}
4	Pólya	0·00026	0·0340	0·0336
	Poisson	0	0·0318	4×10^{-8}
≥ 5	Pólya	0·00013	0·0837	0·764
	Poisson	0	0·0103	1

between the two distributions growing more and more distinct with increasing values of l, especially for the higher values of y_c, i. e. the more energetic showers, the Pólya distribution giving again much higher values for finding small showers (N small). We should, therefore, again expect to find quite different Rossi curves with the Pólya than with the Poisson distribution, but here again the averaging over the primary energy spectrum will smooth out this difference (cf. § 6.4).

§ 6.3. As discussed in § 5.1 we now have to *average over the energy spectrum of the primary photon*, the probability distribution we deduced in § 6.1 for photon-initiated showers. We shall, however not here discuss the deduction of this spectrum as in this paper we are only interested in the fluctuation problem proper. We shall, therefore, only quote the result of A & E who give reasons for the following form of the photon spectrum

$$F(k_0)dk_0 = \begin{cases} c\dfrac{dk_0}{k_0} & \text{for} \quad E' \leq k_0 \leq E_c^{air} \\ c\left(\dfrac{E_c^{air}}{k_0}\right)^{\gamma}\dfrac{dk_0}{k_0} & \text{for} \quad k_0 \geq E_c^{air} \end{cases} \quad . \tag{1}$$

Here $E_c^{air} = 1{\cdot}5 \times 10^8$ e. v. is the critical energy of air (cf. table 11 § 5.2), γ a numerical factor of the order 1–2, which we shall put equal to 1·5, and c the normalization constant determined by the condition

$$\int_{E'}^{\infty} F(k_0)\,dk_0 = 1 \; . \tag{2}$$

Finally, E' is the lower limit of the spectrum which due to the divergence of $\int_0 \frac{dk}{k}$ cannot be put equal to the natural limit 0. We shall take $E' = 10^7$ e. v. The spectrum (1) represents, of course, only a rough estimate. Unfortunately, it cannot be verified by direct experiments, but as discussed in A & E the resulting Rossi curves do not seem to depend significantly on the form of the spectrum (and nor on the value of E').

By means of (1) and the adjusted Pólya distribution obtained in § 6.1[1]) we have, by graphical integration, evaluated the *average* probabilities we are looking for (cf. § 5.1)

[1]) As the Pólya distribution obtained in § 6.1 applies only to *positive* values of
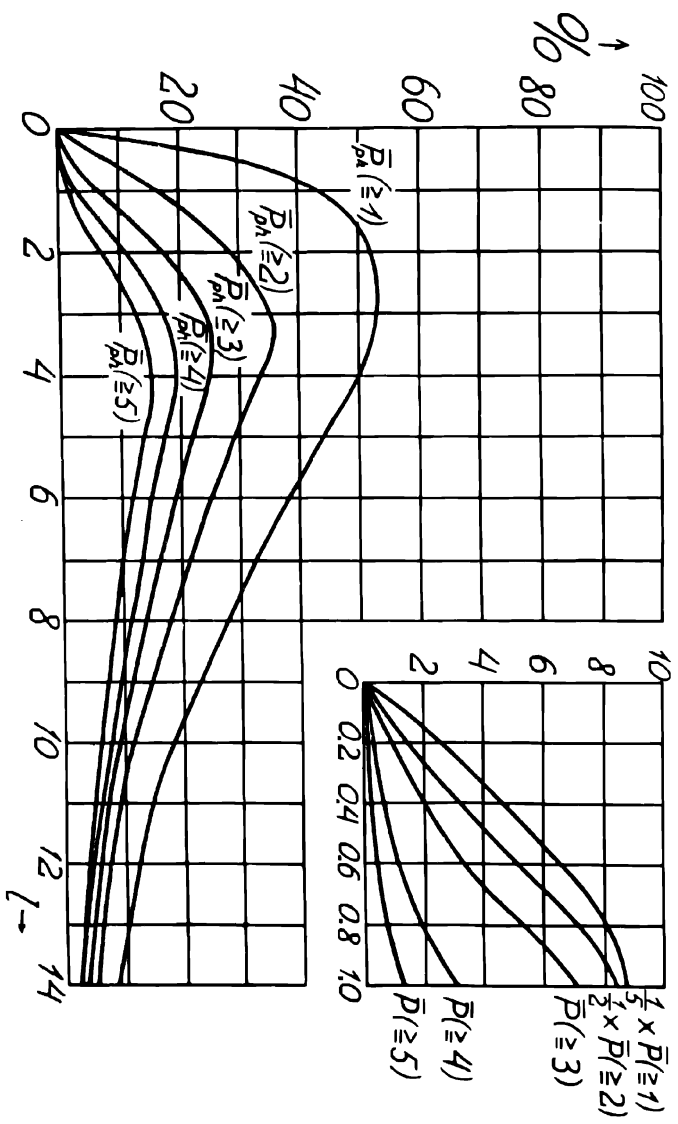
FIG. 27. *The average Pólya distribution* $\bar{P}_{ph}(\geqq N, l)$ *in lead (cf. table 18). The figure in the corner gives the beginning of the curves on a larger scale.*

$$\overline{P}_{ph}(N, l) = \int_{E'}^{\infty} P_{ph}(N, l, k_0) F(k_0) dk_0 . \tag{3}$$

As we are only interested in the small showers measured in the Rossi curves, we have taken $N = 0, 1, 2, 3, 4$ and $N \geq 5$ in which

$$\overline{P}_{ph}(\geq N, l) = \sum_{n=N}^{\infty} \overline{P}_{ph}(n, l) . \tag{4}$$

The result is given (for *lead*) in table 17. The case $N = 0$ has been included as a control of the calculation, as we must have

$$\sum_{n=0}^{\infty} \overline{P}_{ph}(n, l) = 1 . \tag{5}$$

Due to the uncertainties inherent in graphical integration we cannot, however, expect this relation to be strictly fulfilled, but as seen from the last column of table 17 the error is only of the order 10^{-3}–10^{-2}. As discussed at the end of § 5.1 we next have to calculate $\overline{P}_{ph}(\geq N, l)$ as given in (4). The result is given in table 18 and fig. 27. As the *absolute* values cannot be compared with experiments it is interesting to compare the *form* of the curves, arising from the Poisson and the Pólya distributions. In figs. 28–32 we have plotted the two sets of curves together, the Poisson curves being those of A & E, fig. 2. Both curves have been multiplied by a constant factor so as to make the maxima of equal height, the ordinates thus being in arbitrary units.

Looking at the figures we see that, as foreshadowed in § 6.1, *the averaging over the energy spectrum really does smooth out the difference between the Poisson and the Pólya Rossi curves, the difference being in fact very small.* The *form* is practically the same, the difference lying only in the details of the curves and will in most cases be less than the experimental errors. We may e. g. observe that for $\overline{P}(\geq 1, l)$ the Poisson curve is somewhat broader than the Pólya curve, a fact which will make the latter fit better to the experimental curves than the former (cf. § 6.5). We also observe that the difference increases with increasing values of N and that the maxima of the Pólya curves are displaced to the left, i. e. to smaller thicknesses,

the energy parameter y_c, we have quite another probability distribution for *negative* values. With $E' = 10^7$ e. v. no negative values will, however, occur for *lead*, in contrast to other materials. The distribution for negative values of y_c is thus of no interest here and we therefore only refer to the discussion in A & E.
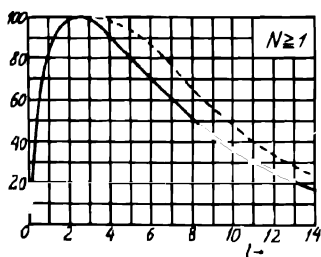
FIG. 28.
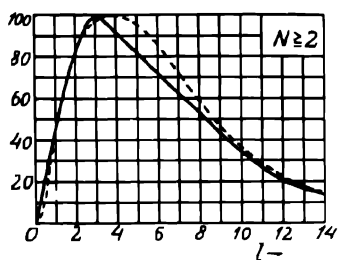
FIG. 29.
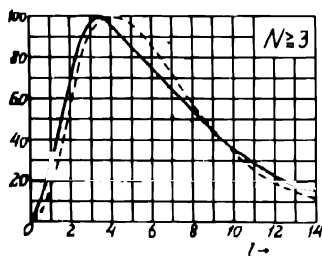
FIG. 30.
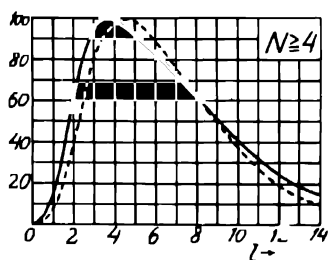
FIG. 31.
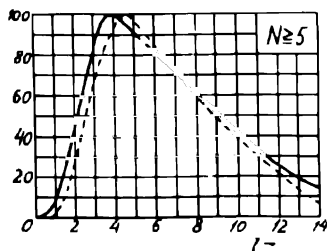
FIG. 32.

FIG. 28–32. *Comparison between the Pólya (full curve) and Poisson (dotted curve) distributions* $\overline{P}_{ph}(\geq N, l)$ *in lead, both sets of curves being multiplied by a constant factor to give a common maximum height of* 100.

Table 17. *Photon-initiated showers in lead. The average probability distribution* $\bar{P}_{ph}(N, l)$ *giving the average probability of a shower containing* ***exactly*** N *electrons as a function of* l.

l \ N		0	1	2	3	4	≧5	$\sum_{n=0}^{\infty} P(n) = 1$
0·2	Pólya	0·865	0·102	0·0191	0·0054[illegible]	0·00245	0·00148	0·995
	Poisson	0·824	0·150	0·0132	0·00069	0·00005	0·000003	0·988
0·4	Pólya	0·757	0·170	0·0426	0·0129	0·004[illegible]	0·00205	0·980
	Poisson	0·71[illegible]	0·248	0·028	0·009	0	0	1·000
0·6	Pólya	0·664	0·227	0·0683	0·0209	0·00711	0·00361	0·991
	Poisson	0·625	0·297	0·061	0·017	0	0	1·000
0·8	Pólya	0·599	0·205	0·0890	0·0331	0·0127	0·00717	1·006
	Poisson	0·554	0·317	0·0092	0·0241	0·00471	0·00059	1·000
1	Pólya	0·551	0·270	0·0975	0·0405	0·0167	0·012[illegible]	0·988
	Poisson	0·485	0·327	0·129	0·0417	0·00987	0·00230	0·995
2	Pólya	0·485	0·235	0·116	0·0620	0·0416	0·0689	1·009
	Poisson	0·380	0·255	0·188	0·101	0·0436	0·0296	0·997
3	Pólya	0·471	0·182	0·114	0·0676	0·0461	0·128	1·009
	Poisson	0·378	0·199	0·150	0·111	0·0714	0·0907	1·000
5	Pólya	0·584	0·145	0·0732	0·0449	0·0322	0·136	1·015
	Poisson	0·415	0·171	0·120	0·0907	0·0610	0·135	1·002
7	Pólya	0·681	0·106	0·0548	0·0345	0·0228	0·109	1·008
	Poisson	0·540	0·145	0·105	0·061	0·049	0·100	1·000
10	Pólya	0·801	0·0693	0·0286	0·0164	0·0117	0·0628	0·990
	Poisson	0·704	0·138	0·055	0·028	0·019	0·056	1·000
14	Pólya	0·898	0·0409	0·0152	0·0096	0·0055	0·0222	0·991
	Poisson	0·852	0·11	0·01	0·01	0·01	0·01	1·002

compared with the Poisson curves. Unfortunately, it is extremely difficult to determine experimentally the Rossi curves for the higher values of N as the intensity, of course, decreases very much with increasing values of N, becoming of the same order as the zero point effect. In § 6.5 we shall compare these results with experiments.

§ 6.4. In this paragraph we give the *averaging over the energy spectrum of the primary electrons*. We shall again only quote the result of A who gives reasons for the following form of the electron spectrum[1])

---

[1]) Cf. also HEITLER (1939).

Table 18. *Photon-initiated showers in lead. The average probability distribution* $\bar{P}_{ph}(\geqq N, l)$ *giving the average probability of a shower containing* **at least** *N electrons as a function of l.*

l \ N		$\geqq 1$	$\geqq 2$	$\geqq 3$	$\geqq 4$	$\geqq 5$
0·2	Pólya	0·130	0·0284	0·00934	0·00393	0·00148
	Poisson	0·164	0·0139	0·00074	0·000053	0·000003
0·4	Pólya	0·232	0·0620	0·0194	0·00645	0·00205
	Poisson	0·285	0·037	0·009	0	0
0·6	Pólya	0·327	0·100	0·0316	0·0107	0·00361
	Poisson	0·375	0·078	0·017	0	0
0·8	Pólya	0·407	0·142	0·0530	0·0199	0·00717
	Poisson	0·446	0·129	0·0297	0·00560	0·00089
1	Pólya	0·437	0·167	0·0699	0·0294	0·0127
	Poisson	0·510	0·183	0·0540	0·0123	0·00239
2	Pólya	0·524	0·289	0·173	0·111	0·0689
	Poisson	0·617	0·362	0·174	0·0732	0·0296
3	Pólya	0·538	0·356	0·242	0·174	0·128
	Poisson	0·622	0·423	0·273	0·162	0·0907
5	Pólya	0·431	0·286	0·213	0·168	0·136
	Poisson	0·587	0·416	0·287	0·196	0·135
7	Pólya	0·327	0·221	0·166	0·132	0·109
	Poisson	0·460	0·315	0·210	0·149	0·100
10	Pólya	0·189	0·120	0·0909	0·0745	0·0628
	Poisson	0·296	0·158	0·103	0·075	0·056
14	Pólya	0·0934	0·0525	0·0373	0·0277	0·0222
	Poisson	0·15	0·04	0·03	0·02	0·01

$$F(E_0)dE_0 = \begin{cases} \alpha \dfrac{dE_0}{E'} & \text{for} \quad 0 \leqq E_0 < E' \\ c\left(\dfrac{E'}{E_0}\right)^{1+\gamma} \dfrac{dE_0}{E'} & \text{for} \quad E_0 \geqq E' \, . \end{cases} \tag{1}$$

Here $E' = 2 \times 10^{8}$ e. v. is an experimentally given energy, α a numerical constant which we shall put equal to 10, $\gamma = 1{\cdot}5$ the same constant as in (6.3.1) and c the normalization constant determined by the condition

$$\int_0^\infty F(E_0)dE_0 = 1 \, . \tag{2}$$

We observe that this spectrum is deduced partly from theoretical arguments and partly from experiments. It is, therefore, somewhat more reliable than the photon spectrum (6.3.1) which was deduced purely theoretically. The experimental verification is in the case of electrons theoretically possible by means of energy measurements in Wilson-chambers, but in practice it is very difficult to perform such measurements without selecting the more energetic electrons and thus disturbing the real statistical distribution. As discussed in A, the Rossi curves do not seem, however, either in this case to depend significantly on the form of the spectrum.

With the spectrum (1) negative values of the energy parameter y_c will occur for *all* materials, in contrast to the photon-initiated showers (cf.[1]) p. 165). However, on our general assumptions, cf. § 5.2, the only particle capable of emerging from the plate in the case $y_c < 0$, i. e. $E_0 < E_c$, is the primary electron itself slowed down by ionization. The negative values, therefore, only contribute to the probabilities of 0 and 1 electron emerging. As shown in A we have, using the Δ-symbol given in (1.2.10),

$$\left.\begin{aligned} P(0, l, E_0) &= \Delta(y_c < \ln(\ln 2 \cdot l)) \\ P(1, l, E_0) &= \Delta(\ln(\ln 2 \cdot l) \leqq y_c \leqq 0) \\ P(\geqq 2, l, E_0) &= 0 \end{aligned}\right\} \text{for } E_0 < E_c, \text{ i. e. } y_c < 0. \tag{3}$$

These formulae simply express the fact (cf. (5.3.1)) that $\beta l \leqq E_0 \leqq E_c$ is just the condition for the primary electron having sufficient energy to penetrate a plate of thickness l without, however, producing a shower.

By means of (1), (3) and the adjusted Pólya distribution (6.2.8) obtained in § 6.2 we have, by graphical integration, evaluated the *average* probabilities we are looking for (cf. § 5.1)

$$\overline{P}_{el}(N, l) = \int_0^\infty P_{el}(N, l, E_0) F(E_0) dE_0 \tag{4}$$

and next

$$\overline{P}_{el}(\geqq N, l) = \sum_{n=N}^{\infty} \overline{P}_{el}(n, l) . \tag{5}$$

As in the case of photon-initiated showers (cf. § 6.3) we have carried out the calculation for $N = 0, 1, 2, 3, 4$ and $N \geqq 5$, the case $N = 0$ being again included in order to make possible the control

$$\sum_{n=0}^{\infty} \overline{P}_{el}(n, l) = 1 . \tag{6}$$

The results are given (for *lead*) in tables 19 and 20 and in fig. 33.
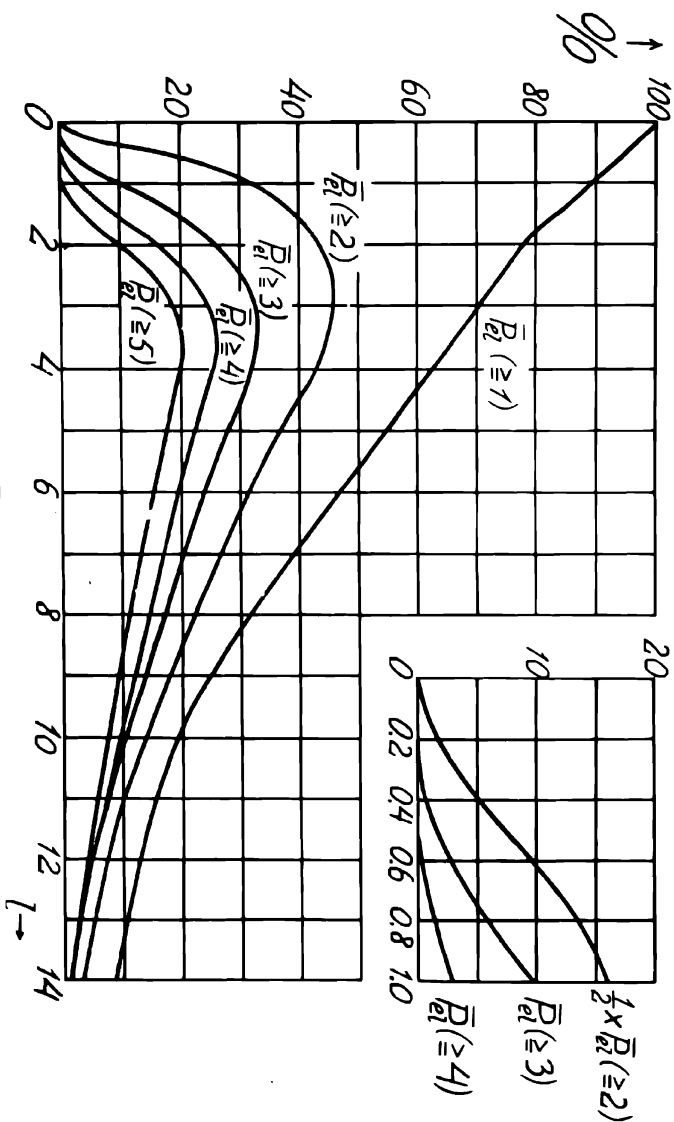
FIG. 33. *The average Pólya distribution $\overline{P}_{el}(\geq N, l)$ in lead (cf. table 20). The figure in the corner gives the beginning of the curves on a larger scale.*

Table 19. *Electron-initiated showers in lead. The average probability distribution* $\overline{P}_{el}(N, l)$ *giving the average probability of a shower containing* **exactly** *N electrons (primary as well as secondary) as a function of l.*

l \ N		0	1	2	3	4	≥5	$\sum_{n=0}^{\infty} P(n) = 1$
0·2	Pólya / Poisson	0·021	0·949	0·0295	0·0005	0	0	1·000
0·4	Pólya / Poisson	0·0398	0·858	0·0979	0·0067	0·00033	0	1·002
0·6	Pólya / Poisson	0·0673	0·740	0·168	0·0256	0·0023	0·0002	1·004
0·8	Pólya	0·085	0·640	0·212	0·0470	0·0100	[illegible]	1·001
	Poisson	0·083	0·625	0·235	0·0470	0·0076	0·0015	0·999
1	Pólya	0·101	0·572	0·222	0·0727	0·0239	0·0030	0·995
	Poisson	0·097	0·525	0·269	0·0797	0·0208	0·0040	0·996
2	Pólya	0·228	0·341	0·174	0·0955	0·0625	0·100	1·001
	Poisson	0·179	0·251	0·234	0·160	0·0817	0·0947	1·000
3	Pólya	0·207	0·245	0·130	0·0802	0·0598	0·188	1·000
	Poisson	0·179	0·211	0·168	0·144	0·103	0·195	1·000
5	Pólya	0·459	0·175	0·0891	0·0590	0·0432	0·175	1·000
	Poisson	0·247	0·179	0·147	0·0934	0·0455	0·288	1·000
7	Pólya	0·607	0·128	0·0583	0·0378	0·0234	0·146	1·000
	Poisson	0·401	0·215	0·0722	0·0225	0·0066	0·283	1·000
10	Pólya	0·814	0·050	0·029	0·014	0·008	0·085	1·000
	Poisson	0·748	0·124	0·029	0·011	0·005	0·083	1·000
14	Pólya	0·920	0·049	0·012	0·007	0·001	0·011	1·000
	Poisson	0·853	0·126	0·010	0·004	0·001	0·006	1·000

As in § 6.3 the relation (6) cannot be fulfilled exactly, but from the last column of table 19 we see that the error is again of the order 10^{-3}–10^{-2}. In order to compare the *form* of the curves arising from the Poisson and the Pólya distributions, we have in figs. 34–38 plotted the two sets of curves together[1]). As in § 6.3 we have again multiplied both curves by a constant factor so as to make the maxima of equal heights, the ordinates thus being in arbitrary units.

[1]) We note that the Poisson curves are not identical with the curves given in A. The curves in the present paper include *all* the 'slow' electrons (cf. § 5.3) whereas in A only $\overline{N}_s$(el) was included.
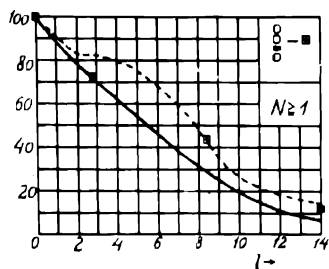
Fig. 34.
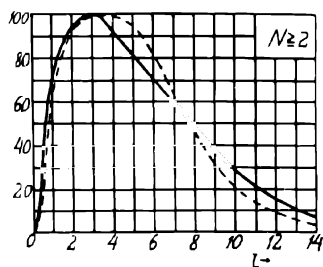
Fig. 35.
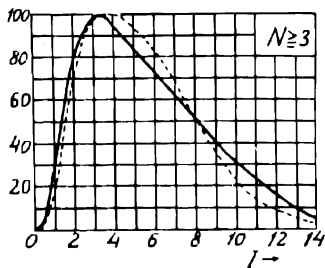
Fig. 36.
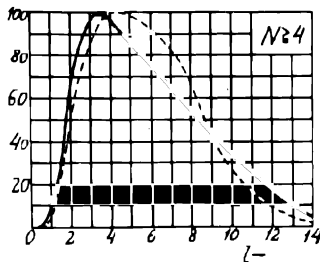
Fig. 37.
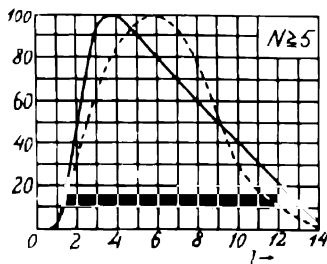
Fig. 38.

Fig. 34–38. *Comparison between the Pólya (full curves) and Poisson (dotted curves) distributions $\overline{P}_{el}(\geq N, l)$ in lead, both sets of curves being multiplied by a constant factor to give a common maximum height of* 100.

Table 20. *Electron-initiated showers in lead. The average probability distribution* $\bar{P}_{el}(\geq N, l)$ *giving the average probability of a shower containing* **at least** N *electrons (primary as well as secondary) as a function of* l.

l \ N		≧1	≧2	≧3	≧4	≧5
0·2	Pólya / Poisson	0·979	0·0295	0·0005	0	0
0·4	Pólya / Poisson	0·963	0·105	0·0070	0·0003	0
0·6	Pólya / Poisson	0·936	0·196	0·028	0·0025	0·0002
0·8	Pólya	0·916	0·270	0·0582	0·0112	0·001[illegible]
	Poisson	0·916	0·291	0·0561	0·0091	0·0015
1	Pólya	0·894	0·322	0·100	0·027	0·003
	Poisson	0·899	0·374	0·105	0·0257	0·0049
2	Pólya	0·773	0·432	0·258	0·163	0·100
	Poisson	0·821	0·570	0·336	0·176	0·0947
3	Pólya	0·703	0·458	0·328	0·248	0·188
	Poisson	0·821	0·610	0·442	0·298	0·195
5	Pólya	0·541	0·366	0·277	0·218	0·175
	Poisson	0·753	0·574	0·427	0·334	0·288
7	Pólya	0·394	0·266	0·207	0·169	0·146
	Poisson	0·599	0·384	0·312	0·290	0·283
10	Pólya	0·186	0·136	0·107	0·093	0·085
	Poisson	0·252	0·128	0·099	0·088	0·083
14	Pólya	0·080	0·031	0·019	0·012	0·011
	Poisson	0·147	0·021	0·011	0·007	0·006

Looking at the figures we see that, as announced in § 6.2, *the averaging over the energy spectrum really does smooth out the difference between the Poisson and the Pólya Rossi curves, the difference being in fact very small*, just as was the case with the photon-initiated showers. Again the difference lies only in the details of the curves and will again in most cases be less than the experimental errors. We first observe that the curve of $\bar{P}(\geq 1, l)$, the *absorption curve*, has a smoother form with the Pólya than with the Poisson distribution. The reason for the curious bump about $l = 2$ is, by the way, that this curve is, so to speak, the superposition of two curves, one

giving the absorption of the primary particle, the other the production of secondaries. We have in the same fig. given the corresponding experimental points of AUGER and collaborators[1]) corresponding to the experimental arrangement indicated in the corner of the fig., but as this experiment is very inaccurate (cf. § 6.5) it cannot distinguish between the two distributions. We next see the same fact as in the case of photon-initiated showers, namely that the difference between the Poisson and the Pólya curves increases with increasing values of N and that the maxima of the Pólya curves are again displaced towards smaller thicknesses compared with the Poisson curves. Again it is, of course, due to the decreasing intensity extremely difficult to determine experimentally the Rossi curves for the higher values of N.

In many experimental arrangements the *total* effect of both photons and electrons is measured, as discussed in § 5.1. We therefore have to calculate a weighted mean of the photon and the electron curves,

$$\overline{P}(\geq N, l) = p_{ph}\overline{P}_{ph}(\geq N, l) + p_{el}\overline{P}_{el}(\geq N, l) . \tag{7}$$

Here p_{ph} and p_{el} denote the probability of a particle of the soft component being a photon and an electron, respectively. Now, the theory is almost symmetric in photons and electrons and the numerical results of the cascade theory therefore also show[2]) that there is approximately an even chance of an incidental particle to be a photon and to be an electron. We thus have to put

$$p_{ph} = p_{el} = \tfrac{1}{2} . \tag{8}$$

From (7), (8) and tables 17–20 we obtain the results given in tables 21 and 22. In figs. 39–43 we give for comparison the Pólya and the Poisson curves together, again adjusted to give the same maximum heights. In fig. 39 we have also given the experimental points of AUGER and collaborators[1]) corresponding to the experimental arrangement indicated in the corner of the fig. In fig. 40 we have given the experimental points of SCHWEGLER[3]) (cf. the discussion in § 6.5). As discussed in the next paragraph both these experiments

[1]) AUGER, LEPRINCE-RINGUET et EHRENFEST Jr. (1936).
[2]) Cf. A & E p. 24.
[3]) SCHWEGLER (1935).
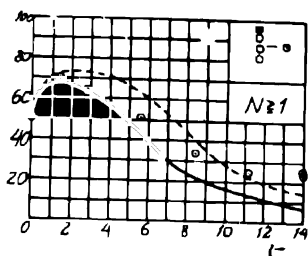
Fig. 39.
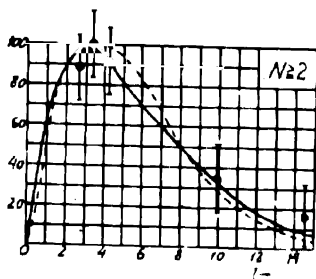
Fig. 40.
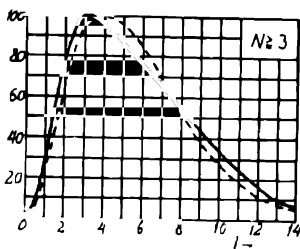
Fig. 41.
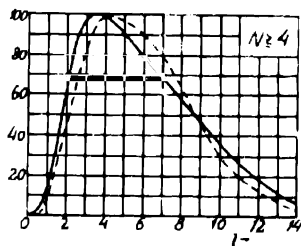
Fig. 42.
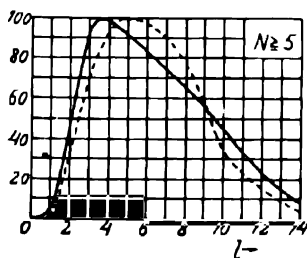
Fig. 43.

Fig. 39–43. *Comparison between the Pólya (full curves) and Poisson (dotted curves) distributions* $\frac{1}{2}\overline{P}_{pk}(\geqq N, l)+\frac{1}{2}\overline{P}_{d}(\geqq N, l)$ *in lead, both sets of curves being multiplied by a constant factor to give a common maximum height of* 100.

Table 21. *Mean value of photon- and electron-initiated showers in lead.* $\frac{1}{2}\bar{P}_{ph}(N, l)+\frac{1}{2}\bar{P}_{el}(N, l)$ *giving the average probability of a shower containing* ***exactly*** *N electrons, as a function of l, when the primary particle initiating the shower has the probability* $\frac{1}{2}$ *of being a photon and* $\frac{1}{2}$ *of being an electron.*

l \ N		0	1	2	3	4	≧5	$\sum_{n=0}^{\infty} P(n) = 1$
0·2	Pólya	0·443	0·526	0·0243	0·0030	0·0012	0·001	0·999
	Poisson	0·423	0·550	0·0214	0·0006	0·00003	0·000002	0·995
0·4	Pólya	0·398	0·514	0·0703	0·0098	0·0024	0·0010	0·996
	Poisson	0·377	0·553	0·063	0·008	0·0002	0	1·001
0·6	Pólya	0·366	0·484	0·118	0·0233	0·0047	0·0019	0·998
	Poisson	0·346	0·519	0·115	0·021	0·001	0·0001	1·002
0·8	Pólya	0·342	0·456	0·151	0·0401	0·0114	0·0042	1·005
	Poisson	0·319	0·471	0·167	0·0356	0·0062	0·0012	1·000
1	Pólya	0·326	0·421	0·160	0·0566	0·0203	0·0079	0·992
	Poisson	0·291	0·426	0·199	0·0607	0·0153	0·0036	0·996
2	Pólya	0·357	0·288	0·145	0·0792	0·0521	0·084	1·005
	Poisson	0·280	0·253	0·211	0·131	0·0627	0·0622	1·000
3	Pólya	0·384	0·214	0·122	0·0739	0·0530	0·158	1·005
	Poisson	0·279	0·205	0·159	0·128	0·087	0·143	1·001
5	Pólya	0·522	0·160	0·0812	0·0520	0·0377	0·156	1·009
	Poisson	0·331	0·175	0·138	0·092	0·053	0·212	1·001
7	Pólya	0·644	0·117	0·0566	0·0362	0·0231	0·128	1·005
	Poisson	0·471	0·180	0·089	0·042	0·028	0·192	1·002
10	Pólya	0·808	0·060	0·029	0·015	0·010	0·074	0·996
	Poisson	0·726	0·131	0·042	0·020	0·012	0·070	1·001
14	Pólya	0·909	0·045	0·14	0·009	0·003	0·017	0·997
	Poisson	0·853	0·12	0·01	0·007	0·006	0·008	1·004

are, however, as also shown by the large mean errors, far too inaccurate to decide between the two distributions.

§ 6.5. We shall now finally compare our results with experiments. This comparison is, however, in fact rather difficult of performance as only few experiments among the many investigations of cosmic rays are sufficiently refined for this purpose. As previously mentioned (cf. § 5.1) most Rossi curves include the *total* effects of both

Table 22. *Mean value of photon- and electron-initiated showers in lead.* $\frac{1}{2}\bar{P}_{ph}(\geq N, l)+\frac{1}{2}\bar{P}_{el}(\geq N, l)$ *giving the average probability of a shower containing* **at least** *N electrons, as a function of l, when the primary particle initiating the shower has the probability* $\frac{1}{2}$ *of being a photon and* $\frac{1}{2}$ *of being an electron.*

l \ N		≥1	≥2	≥3	≥4	≥5
0·2	Pólya	0·555	0·0295	0·0052	0·002	0·001
	Poisson	0·572	0·0220	0·0006	0·00003	0·000002
0·4	Pólya	0·508	0·0835	0·0132	0·0034	0·0010
	Poisson	0·624	0·0712	0·0082	0·0002	0
0·6	Pólya	0·632	0·148	0·0299	0·0066	0·0019
	Poisson	0·656	0·137	0·0221	0·0011	0·0001
0·8	Pólya	0·663	0·207	0·0557	0·0156	0·0042
	Poisson	0·681	0·210	0·0430	0·0074	0·0012
1	Pólya	0·666	0·245	0·0848	0·0282	0·0079
	Poisson	0·705	0·279	0·0796	0·0189	0·0036
2	Pólya	0·648	0·360	0·215	0·136	0·084
	Poisson	0·720	0·467	0·256	0·125	0·0622
3	Pólya	0·621	0·407	0·285	0·211	0·158
	Poisson	0·722	0·517	0·358	0·230	0·143
5	Pólya	0·487	0·327	0·246	0·194	0·156
	Poisson	0·670	0·495	0·357	0·265	0·212
7	Pólya	0·361	0·244	0·187	0·151	0·128
	Poisson	0·531	0·351	0·262	0·220	0·192
10	Pólya	0·188	0·128	0·099	0·084	0·074
	Poisson	0·275	0·144	0·102	0·082	0·070
14	Pólya	0·088	0·043	0·029	0·020	0·017
	Poisson	0·151	0·031	0·021	0·014	0·008

primary electrons, photons ('soft' showers) and mesons ('hard' showers). Before we can compare the theoretical curves with the corresponding experimental curves we have, therefore, to *separate the different components experimentally, but although this seems a natural task only few experiments have so far been carried out on such lines.*

We shall begin our discussion with the experiment of AUGER and collaborators[1]) already referred to in the preceding paragraph. This

[1]) AUGER, LEPRINCE-RINGUET et EHRENFEST Jr. (1936).

experiment isolates in a certain way the effects of the soft component from those of the hard one. They measure triple coincidences between 3 counters placed in a vertical line. In the first arrangement the shower-producing material was placed *between* the two lowest counters as indicated in the corner of fig. 34 p. 174. They thus measure the *absorption curve* of the ionizing particles and the result is shown in fig. 44. Extrapolating the curves backwards from large
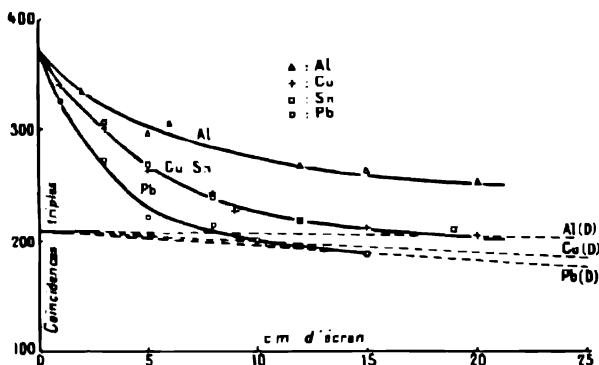
FIG. 44. *The experimental results of Auger obtained with the arrangement indicated in the corner of fig. 34. The dotted curves give the curves extrapolated backwards from large thicknesses.*

thicknesses to small thicknesses and subtracting these extrapolated curves from the total curves they obtain that part of the absorption curves which is due to the soft component, i.e. $\overline{P}_{el}(\geq 1, l)$. In fig. 34 p. 174 we have compared AUGER's points with the theoretical curves for lead[1]). *The agreement, though satisfactory, shows little* in as much as a subtraction between two independent curves, one of which is even obtained by extrapolation, will of course entail very large statistical errors. In a second arrangement AUGER placed the shower-producing material *above* all the counters (cf. fig. 39 p. 177), thus measuring a sort of absorption curve of the whole incidental radiation. Again extrapolating backwards from large thicknesses and subtracting they obtain points which may be compared with the theoretical $\frac{1}{2}\overline{P}_{ph}(\geq 1, l) + \frac{1}{2}\overline{P}_{el}(\geq 1, l)$ curve. In fig. 39 we have compared AUGER's points with the theoretical curves. *The same remarks apply to this comparison as to the previous one.*

[1]) For the corresponding Poisson curve for aluminium, cf. A p. 537.

The first experiment in which a direct separation of the soft and hard showers was aimed at is the experiment of SCHWEGLER[1]), in which were measured triple coincidences from a lead plate, P, using the arrangement shown in fig. 45. It is seen, firstly, that the showers measured can be produced *both* by primary ionizing and
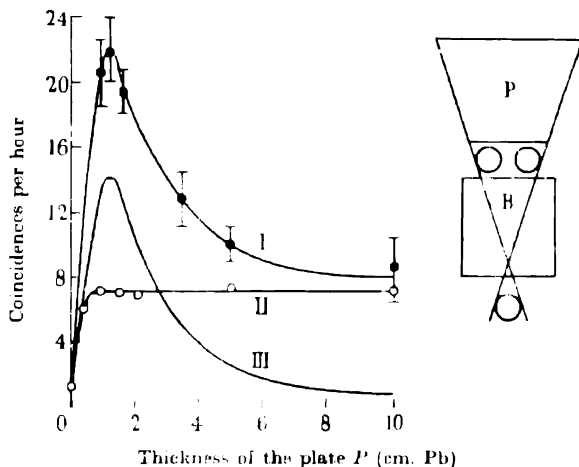
FIG. 45. *The experimental arrangement and results of Schwegler.*

non-ionizing rays and, secondly that they must contain at least *two* ionizing particles. Curve I is obtained *without* the 10 cm thick lead block, B, curve II *with* B placed between the counters. Curve III, being the difference between I and II, is then supposed to give the effects due to the soft component alone. It is, however, on the one hand obvious that in this arrangement we do *not* catch *all* the hard showers, as the paths of the mesons, producing showers, need not all lie within the cone determined by the counters. E. g. the shower shown in fig. 46a is included in I, but not in II. Consequently, III is *too high*, as it may contain some meson-produced showers. We ought, therefore, place a whole floor of counters extending to both sides below B in order to catch as many mesons as possible. On the other hand, a meson passing through B may eject one or more secondary electrons and thus give rise to a triple coincidence which need not have been counted without B. E. g. the shower in fig. 46b

[1]) SCHWEGLER (1935).

is included in II, but not in I. Consequently, III is *too low* as we may have subtracted too many meson-produced showers. To avoid the last mentioned source of error the separating filter B ought to be placed *below* the system of coincidence counters. Finally, it is evident that the statistical measuring errors will be very large due
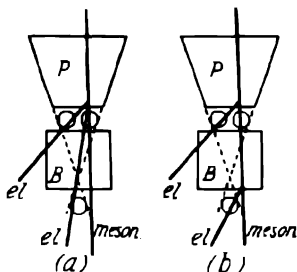
FIG. 46. *Two sources of error in Schweglers experiment (cf. the text)*

to this method of subtraction. Consequently, nearly any theoretical curve may be fitted to these points as may be seen from fig. 40 p. 177. We have here compared SCHWEGLER's points (from curve III) with the theoretical $\frac{1}{2}\overline{P}_{ph}(\geq 2, l)+\frac{1}{2}\overline{P}_{el}(\geq 2, l)$ curves. As was the case with AUGER's experiment *the agreement, though excellent, is not conclusive*.

In A & E[1]) the author has suggested the application of the method of *anti-coincidences*, originally employed by ROSSI to isolate the effects of the photons (cf. below), for the separation of the soft and hard showers. Experiments on such lines are now being carried out by AMBROSEN. But as these experiments have not yet been concluded, it is too early to discuss his results.

As just mentioned the method of anti-coincidences was first employed for isolating the showers initiated by photons only. The experimental arrangement used by ROSSI and JÁNOSSY[2]) for this purpose is shown in fig. 47. Here s is the shower-producing lead plate. The five upper counters A are coupled in parallel, as are also the two lower counters D. By this arrangement they measure *simultaneously* the number of coincidences B-C-D (n_1) and A-B-C-D (n_2). The difference n_1-n_2, consequently, gives the

[1]) Cf. A & E p. 38.
[2]) ROSSI and JÁNOSSY (1939a).

number of *anti-coincidences*, i. e. of those coincidences B-C-D which are *not* accompanied by a simultaneous discharge of the counter battery A. As both terms in the difference are measured simultaneously the statistical measuring errors will again be very small. As a non-ionizing particle does not discharge a counter these anti-
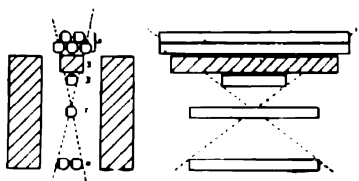
FIG. 47. *The experimental arrangement of Rossi and Jánossy.*

coincidences must be due to *non-ionizing* primary particles, i. e. either photons, neutral mesons or neutrons, producing showers containing at least one ionizing particle. Primary ionizing rays would, obviously, discharge A and could thus not give rise to an anti-coincidence. Now the neutrons may be excluded at once, partly because there are presumably only very few neutrons in cosmic radiation and partly because they would, probably, have very little chance of producing secondary ionizing particles. This experiment may, therefore, elucidate the question of the existence of *neutral* mesons. This question is, as already mentioned in the introduction, of vital importance to the theory of the meson. In fig. 48 we give the experimental results of ROSSI and JÁNOSSY (the small zero-point effect having been subtracted) together with the corresponding theoretical curves $\overline{P}_{\mu h}(\geqq 1, l)$ adjusted by means of the method of least squares. It will, firstly, be seen that the statistical errors are in fact very small. Next, it will be seen that *the agreement is even better than might be expected* and that *the experimental points are in favour of the Pólya distribution*, the difference between the two distributions being larger than the experimental errors.

It is an interesting question whether the fact, found in § 6.1, that the fluctuation is of varying size, is important or whether we might also obtain a good agreement taking the Pólya parameter $b =$ const. Especially, it is interesting to put $b = 1$ as the resulting distribution is that proposed by FURRY (cf. § 4.3 and (4.4.5)). The resulting curve is given by the stippled curve in fig. 48 and it is
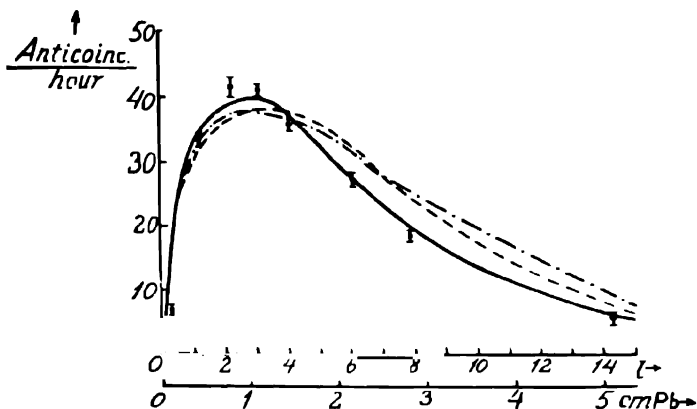
FIG. 48. *The theoretical $\overline{P}_{ph}(\geq 1, l)$ curves fitted by the method of least squares to the experimental points of Rossi and Jánossy. Full, dotted and stippled curve corresponds to the Pólya, Poisson and Furry distribution, respectively.*

seen that the agreement is only poor. *The fluctuation being of varying size is, consequently, essential.*

Finally, we observe that the ROSSI and JÁNOSSY experiment gives a negative answer to the problem of the existence of neutral mesons, as the theoretical curve fits the experimental points even for thicknesses as large as $l = 14$, i. e. 5 cm lead. *This means, that if the hard component contains neutral mesons at all, the probability for these particles to produce ionizing rays must be extremely small.* Neutral mesons would, namely, as discussed above, also give rise to anti-coincidences if they produced ionizing secondaries, and as they would certainly be very penetrating they would contribute mainly to the tail of the curve. Thus, the form of the curve would be quite different. This negative result is, by the way, also in accordance with other experiments. We shall here only mention those of LOVELL[1]) and TRUMPY and BJORDAL[2]). If it is true, as now seems to be the case, that also the *ionizing* mesons have extremely small probabilities of producing ionizing secondary mesons (cf. e. g. the experiment of LOVELL just quoted) *the negative results of the experiments mentioned are just what might be expected theoretically*, as the two probabilities in question must be of the same order of magnitude.

[1]) LOVELL (1939).

[2]) TRUMPY and BJORDAL (1942).

We may note that the agreement between theory and ROSSI and JÁNOSSY's experiment depends, of course, essentially on the form of the primary energy spectrum, and *the agreement may thus be taken as an indirect proof of the validity of the form applied*. For a discussion of the influence of the spectrum we refer to A & E[1]). In this connection we may, however, mention that due to the *backward effect*, i. e. the effect of secondary shower electrons being ejected in the backward direction, it is possible that not all the photons which hit the plate give rise to anti-coincidences. This backward effect would, presumably, be the more pronounced the more energetic the photons. We should, thus, perhaps expect the anti-coincidences to be produced mainly by the less energetic photons, i. e. by an effective spectrum displaced somewhat towards smaller energies, having, e. g., a somewhat higher value of the exponent γ.

The experiment of ROSSI and JÁNOSSY has recently been extended by NERESON[2]) in order to investigate also the showers containing *at least two* and *at least three* electrons. His experimental arrangement is shown in fig. 49. Here s is the shower-producing lead plate and
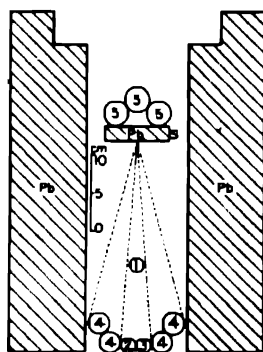
FIG. 49. *The experimental arrangement of Nereson.*

the counters denoted with the same number are coupled in parallel. By an appropriate electrical circuit NERESON measures *simultaneously* the following four quantities:

(1) N_{12345}: the number of coincidences between all the counters 1-5,

(2) N_{1235-4}: the number of coincidences between counters 1, 2, 3 and 5 which are *not* accompanied by a simultaneous discharge of the counters 4,

---

[1]) Cf. A & E § 6.

[2]) NERESON (1942).

(3) N_{1234-5}: the number of coincidences between counters 1, 2, 3 and 4 which are *not* accompanied by a simultaneous discharge of the counters 5,

(4) N_{123-45}: the number of coincidences between counters 1, 2 and 3 which are *not* accompanied by a simultaneous discharge of both counters 4 and 5.

N_{12345} is now interpreted as measuring electron-initiated showers containing at least three electrons, i.e. $\overline{P}_{el}(\geq 3)$, N_{1235-4} as $\overline{P}_{el}(= 2)$, N_{1234-5} as $\overline{P}_{ph}(\geq 3)$ and N_{123-45} as $\overline{P}_{ph}(= 2)$. This interpretation is, however, not correct. Firstly, N_{12345} and N_{1235-4} contain also the showers produced by *mesons* and, secondly, N_{1235-4} and N_{123-45} may also contain some showers with *more* than two ionizing particles, if they are only so narrow, i.e. so energetic, that all the particles are contained within the solid angle of the counters 1, 2 and 3. We can, consequently, only compare NERESON's $N_{1234-5}+N_{123-45}$ curve with our $\overline{P}_{ph}(\geq 2)$ curve, and his N_{1234-5} curve with our $\overline{P}_{ph}(\geq 3)$ curve. In figs. 50 and 51 we give this comparison, the
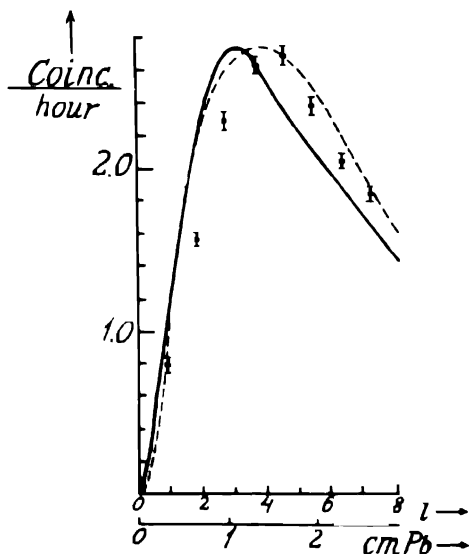
FIG. 50. *The theoretical $\overline{P}_{ph}(\geq 2, l)$ curves fitted to give the same maximum height as the experimental points of Nereson (his $N_{1234-5}+N_{123-45}$ curve, cf. the text). Full and dotted curve corresponds to the Pólya and Poisson distribution, respectively.*
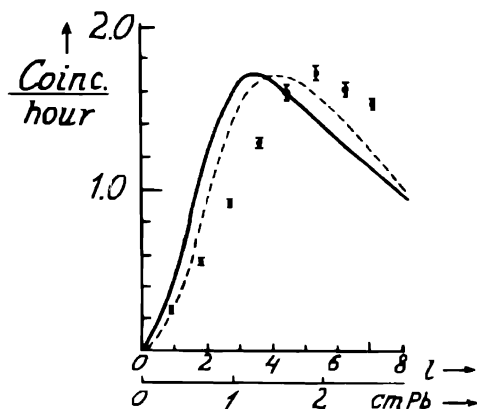
FIG. 51. *The theoretical* $\overline{P}_{ph}(\geq 3, l)$ *curves fitted to give the same maximum height as the experimental points of Nereson (his* N_{1234-4} *curve, cf. the text). Full and dotted curve corresponds to the Pólya and Poisson distribution, respectively.*

theoretical curves being adjusted to give the same maximum heights as the experimental curves. (The zero-point effect which, of course, is quite large in such experiments, viz. about 2 counts per hour, has been subtracted.) As will be seen, the agreement is, both as regards the Poisson and the Pólya curves, not as good as in the case of ROSSI and JÁNOSSY's experiment, i.e. for the $\overline{P}_{ph}(\geq 1)$ curve. It will, however, be noticed, firstly that the experimental maxima are displaced somewhat to the *right* of the theoretical ones and, secondly, that this displacement is larger for the $N \geq 3$ curve than for the $N \geq 2$ curve. *This fact is just what might be expected, as it is quite possible that the experimental arrangement is in favour of the larger showers, and the more so the larger the minimum value, N, of the number of electrons in the showers measured. We therefore ought not take all the geometrical factors equal to one* (*cf.* § 5.1) *but give the higher N-curves higher weights which fact would actually explain the displacement of the experimental maxima towards larger thicknesses*[1]). We shall not here discuss the effect of such a weighing of the various $\overline{P}(N)$ curves as, unfortunately, NERESON has not measured the tails of the curves. We shall, therefore, postpone such a discussion until AMBROSEN has concluded the experiments mentioned

[1]) The same effect has also been observed by AUGER as mentioned in a conversation with the author, but these experiments do not seem to have been published.

above. Finally we observe that the figures seem to show that the Pólya curves will also in the case $N \geqq 2$ and $N \geqq 3$ give a better fit than the Poisson curves, the latter curves being too broad as compared with the experimental curves.

We shall finally discuss the latter experiment separating the soft and the hard showers, namely that of TRUMPY[1]). It is the primary purpose of this experiment to clarify the only point on which there was strong disagreement between theory and experiment, viz. the dependence of the beginning of the Rossi curves on the nature of the shower-producing material, i. e. of Z. The theoretical curves lie, for each fixed N, the higher the larger the value of Z because *larger* Z means *lower* critical energy (cf. (5.2.1)) and, consequently, *larger* relative energies, $y_c = \ln \frac{E_0}{E_c}$, which are the only energies entering into the shower theory (cf. chap. 5). This fact was, however, in contradiction to the experiments which found that different materials give Rossi curves lying on the *same* curve when plotted on an l-scale[2]). As discussed in A and A & E the reason for this discrepancy was conjectured to lie in the fact that the effects of the hard component, not being subtracted in the experiments in question, are of importance down to very small thicknesses. This hypothesis has been verified most beautifully in the recent experiment of TRUMPY, who measures the $\overline{P}_{ph}(\geqq 1, l)$ curves for Pb, Fe and Al with the anti-coincidence arrangement shown in fig. 52. F is the shower producing layer, A a battery of eleven counters coupled in parallel and K_3 another battery of three counters also coupled in parallel. TRUMPY now measures *simultaneously* the numbers of coincidences K_1-K_2-K_3 (n_1) and A-K_1-K_2-K_3 (n_2). The difference $n_1 - n_2$, consequently, gives the number of anti-coincidences, i. e. of those coincidences K_1-K_2-K_3 which are *not* accompanied by a simultaneous discharge of the counter battery A. It will be seen that this arrangement is very nearly the same as that employed by ROSSI and JÁNOSSY as discussed above, and the same discussion therefore applies to the former as to the latter. In fig. 52 we also show the experimental points of TRUMPY and it will be seen that the variation with Z is actually that predicted by the cascade theory. Comparing with the experimental points of

[1]) TRUMPY (1943). I wish to thank Prof. TRUMPY for kindly giving me an opportunity to see his manuscript before publication.

[2]) Cf. A & E § 7.
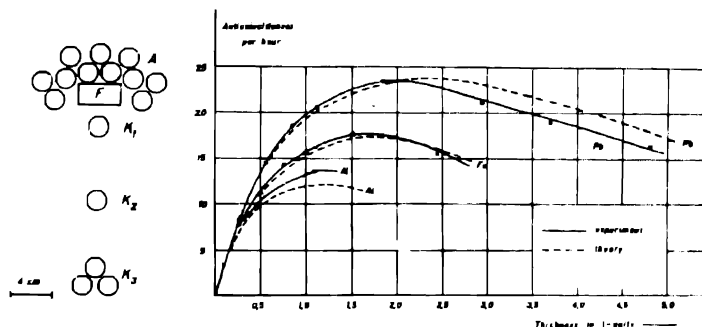
Fig. 52. *The experimental arrangement and results of Trumpy.*

Rossi and Jánossy (cf. fig. 48 p. 184) we see, however, that the Pb curve of Trumpy is displaced towards *smaller* thicknesses. The reason for this fact is, presumably, that the counter walls employed by Trumpy are thicker than those of Rossi and Jánossy, some *'slow' electrons* (*cf.* § 5.3) *thus being unable to penetrate all the counters*. In fact, as seen on fig. 53 the theoretical Pólya $\overline{P}_{ph}(\geq 1, l)$ 'fast' + 'slow' curve, which fitted the points of Rossi and Jánossy so excellently, does *not* fit the points of Trumpy nearly as well. As Trumpy himself estimates that he cannot count shower
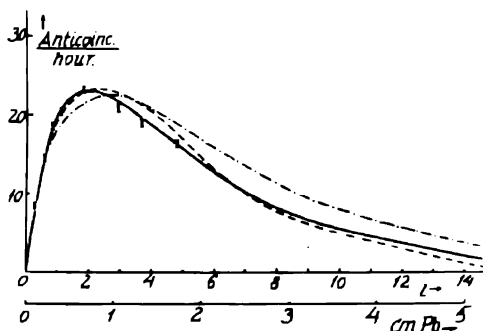
Fig. 53. *The theoretical $\overline{P}_{ph}(\geq 1, l)$ curves fitted by the method of least squares to the experimental lead points of Trumpy. Full and stippled curve corresponds to the Pólya distribution with only 'fast' electrons (i. e. with energy > 10^7 e. v.) and 'fast' + 'slow' electrons included, respectively. Dotted curve corresponds to the Poisson distribution with only 'fast' electrons included.*

electrons with energies *below*, say, 10^7 e. v., which energy happens to coincide with the critical energy of lead (cf. table 11 § 5.2), we have to omit *all* the 'slow' electrons in the case of Pb and recalculate the $\overline{P}_{ph}(\geqq 1, l)$ curves[1]). In fig. 53 we show the resulting Pólya and Poisson[2]) 'fast' curves, adjusted by means of the method of least squares. Although the difference between the Pólya and the Poisson curves is here very small, *the experimental points are definitely in favour of the former*, just as was the case with the experimental points of Rossi and Jánossy. Moreover, *the agreement is again seen to be even better than might be expected.*

Unfortunately, Trumpy has not measured the curve for so high thicknesses as have Rossi and Jánossy but we hope that the tail of the curve will be obtained in further experiments, both for Pb, Fe *and* Al. *We also hope that it may be possible to measure the curves for the larger values of N, again both for* Pb, Fe *and* Al *in order to compare both the variation with N and with Z with experiments. Perhaps it may also be possible to refine the method of anti-coincidences still more so as to allow a separate measurement of the electron-initiated showers. Such an experiment would be of great importance as the primary electron spectrum is better founded than the primary photon spectrum. Finally, we may mention that it is perhaps not sufficient to study the shower phenomenon by means of counters only, as such a study cannot tell us what really happens in the single acts. The ideal procedure would be to enclose the whole counter arrangement in a Wilson-chamber, but whether this is possible we must leave it to the experimentators to decide!*

Concluding this discussion we may summarize that the agreement between theory and experiments is in most cases excellent and that the experiments in fact seem to be in favour of the Pólya distribution suggested in the present paper, as we would have expected it on the basis of our theory. The present experimental material is, however, insufficient for such a detailed discussion of the comparison between theory and experiments as theory permits. Further experiments are, therefore, still highly desirable.

---

1) We observe that it is not sufficient simply to carry out this calculation with the corresponding smaller mean values, but that the fluctuation parameter, b, of the Pólya distribution has also to be re-adjusted. As b depends only on the *relative* form of the l and t curves (cf. § 6.1) this re-adjustment was expected not to influence significantly the magnitude of b, as was actually found to be the case by the numerical calculations.

2) The Poisson curve is simply that given in A & E fig. 12.

PART III

APPENDICES

CHAPTER 7.

On Exponentiable Matrices of one and more Dimensions.

§ 7.1. By an infinite matrix $\boldsymbol{A}$ we understand a matrix with enumerable infinite rows and columns. The elements, which we shall denote by[1] $(i|\boldsymbol{A}|j)$ $i, j = 0, 1, 2, \ldots$, may be real or complex constants or functions of one — or more — variable x, which may also be real or complex[2]. An arbitrary matrix $\boldsymbol{A}$ we shall call **exponentiable**[3] if, firstly, all the powers

$$\boldsymbol{A}^{\nu} \quad \nu = 0, 1, 2, \ldots \quad (\boldsymbol{A}^0 = \boldsymbol{1}) \tag{1}$$

exist, and, secondly, the matrix function

$$\exp \boldsymbol{A} = \sum_{\nu=0}^{\infty} \boldsymbol{A}^{\nu} \frac{1}{\nu!} \tag{2}$$

exists. The product

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C} \tag{3}$$

---

[1]) This symbolism was first introduced by DIRAC. We have chosen it here because it is especially convenient and symmetric and may easily be extended to multi-dimensional matrices, cf. § 7.5.

[2]) Matrices we denote by — small or capital — clarendon letters, $\boldsymbol{A} = \{(i|\boldsymbol{A}|j)\}$ etc. By the *main diagonal* of a matrix we understand the elements $(i|\boldsymbol{A}|i)$ $i = 0, 1, 2, \ldots$. By the matrix $\boldsymbol{1}$ we understand the infinite unit-matrix having the elements $(i|\boldsymbol{1}|j) = \delta_{ij} = \begin{cases} 1 \text{ for } i = j \\ 0 \;\; ,, \;\; i \neq j \end{cases}$. By the matrix $|\boldsymbol{A}|$ we understand the matrix having the elements $(i||\boldsymbol{A}||j) = |(i|\boldsymbol{A}|j)|$. By the inequality $|\boldsymbol{A}| \leq \boldsymbol{B}$ we symbolize that $(i||\boldsymbol{A}||j) \leq (i|\boldsymbol{B}|j)$ for all $i, j = 0, 1, 2, \ldots$. We note that $|\boldsymbol{A} \cdot \boldsymbol{B}| \leq |\boldsymbol{A}| \cdot |\boldsymbol{B}|$ where the product is defined in the usual way given in (7.1.4). Furthermore, we define $\int \boldsymbol{A} dx$ as the matrix having the elements $\int (i|\boldsymbol{A}|j) dx$, $\frac{\partial}{\partial x} \boldsymbol{A}$ as the matrix having the elements $\frac{\partial}{\partial x} (i|\boldsymbol{A}|j)$, $\max\limits_{x_0 \leq t \leq x} \boldsymbol{A}(t)$ as the matrix having the elements $\max\limits_{x_0 \leq t \leq x} (i|\boldsymbol{A}(t)|j)$, and so on.

[3]) In § 7.6 we shall summarize all our definitions and theorems.

is here defined in the usual way as the row-column product

$$(p|\boldsymbol{C}|q) = \sum_{\alpha=0}^{\infty} (p|\boldsymbol{A}|\alpha)(\alpha|\boldsymbol{B}|q) \qquad \text{for all } p, q = 0, 1, 2, \ldots . \qquad (4)$$

An arbitrary matrix we shall call **absolutely exponentiable** if (1) and (2) are also true for $|\boldsymbol{A}|$. Now these notions of exponentiability, which are natural from a purely algebraic point of view, are not very suitable for our purpose before we have modified them somewhat. Let $\boldsymbol{A}$ be a function of a variable x, which function we shall in the following always tacitly assume to be *continuous*. Then the matrix

$$\boldsymbol{K} = \max_{x_0 \leq t \leq x} |\boldsymbol{A}(t)| \qquad (5)$$

exists, as $\boldsymbol{A}$ is continuous in a closed interval. We shall now call **$\boldsymbol{A}$ absolutely exponentiable in the interval $x_0 \leq t \leq x$** if the matrix function $\boldsymbol{K}(x-x_0)$ is exponentiable, i. e. if, firstly, all the powers

$$\boldsymbol{K}^\nu \quad \nu = 0, 1, 2, \ldots \qquad (6)$$

exist and, secondly, the matrix function

$$\exp\left[\boldsymbol{K}(x-x_0)\right] = \sum_{\nu=0}^{\infty} \boldsymbol{K}^\nu \frac{(x-x_0)^\nu}{\nu!} \qquad (7)$$

exists. It is exclusively in this modified form we shall use the notion of exponentiability and we shall, therefore, often simply say that $\boldsymbol{A}$ is exponentiable when (7) is fulfilled.

It is beyond the scope of the present paper to discuss the general conditions for a matrix to be exponentiable[1]). We shall confine ourselves to give some *sufficient* criteria which will, however, cover most cases met with in the practical applications of the theory of stochastic processes.

We firstly observe that due to (2.1.6) and (2.1.12) all the operator matrices met with in the theory of probability have *absolutely convergent column sums* namely, using the symbol introduced in (2.1.10),

$$0 \leq (\boldsymbol{\Sigma} \cdot |\boldsymbol{A}||q) = (\boldsymbol{\Sigma} \cdot (\boldsymbol{\Pi}+\boldsymbol{1}) \cdot \boldsymbol{p}|q) = 2p(q, t) < \infty \qquad (8)$$
$$\text{for all } q = 0, 1, 2, \ldots .$$

We shall, consequently, in the following always assume $\boldsymbol{A}$ to have

---

[1]) The knowledge hitherto obtained in the theory of infinite matrices even indicate that it is impossible to give simple, necessary and sufficient conditions for this property.

absolutely convergent column sums. We note, however, that this condition is not by far sufficient to ensure A being exponentiable as shown by the following example.

Example 1.

Let $\varepsilon > 0$ be arbitrary and let

$$(p|A|q) = p^{-1-\varepsilon} q^{\varepsilon} \qquad \text{for all } p, q = 1, 2, 3, \ldots \quad (\varepsilon > 0) . \tag{9}$$

We then have

$$(\Sigma \cdot A|q) = \sum_{i=1}^{\infty} i^{-1-\varepsilon} q^{\varepsilon} = \zeta(1+\varepsilon) q^{\varepsilon} < \infty \qquad \text{for all } q = 1, 2, 3, \ldots \quad ,, \ ,, \ \varepsilon > 0 \tag{10}$$

where $\zeta(z)$ is the Zeta-function of RIEMANN. A has thus absolutely convergent column sums but even the square of A does not exist because

$$(p|A^2|q) = \sum_{\alpha=1}^{\infty} p^{-1-\varepsilon} \alpha^{\varepsilon} \alpha^{-1-\varepsilon} q^{\varepsilon} = p^{-1-\varepsilon} q^{\varepsilon} \zeta(1) = \infty \qquad \text{for all } p, q = 1, 2, 3, \ldots \quad ,, \ ,, \ \varepsilon > 0 . \tag{11}$$

We observe that the column sums (10) increase only with q as q^{ε} where $\varepsilon (> 0)$ can be arbitrarily small. Example 1 shows, consequently, that *we cannot hope for a general criterium of exponentiability simply to consist in a condition limiting the speed of increase of the absolute column sums*, unless all these sums are simply *uniformly bounded*. Such a matrix we shall call **bounded**[1]), i. e.

$$0 \leqq (\Sigma \cdot |A||q) \leqq M < \infty \qquad \text{for all } q = 0, 1, 2, \ldots \tag{12}$$

where M is a constant.

Lemma 1.

For a bounded matrix all the powers $|A|^{\nu}$ *and, consequently, also* A^{ν}, *exist and are again bounded matrices with the bound* M^{ν}, i. e.

$$|(\Sigma \cdot A^{\nu}|q)| \leqq (\Sigma \cdot |A|^{\nu}|q) \leqq M^{\nu} < \infty \qquad \text{for all } q = 0, 1, 2, \ldots . \tag{13}$$

---

[1]) We note that we here use this term in another sense than that used in the theory of HILBERT space. In that theory an infinite matrix is called bounded if $\left|\sum_{i=1}^{n} \sum_{j=1}^{n} x_i (i|A|j) y_j\right| \leq M \sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{j=1}^{n} y_j^2}$ for all $n = 1, 2, 3, \ldots$. It is obvious that these two definitions are not identical. We ought to use the expression *uniformly column sum bounded*, but no confusion being possible, as the problems dealt with here are quite different, we shall abbreviate it to *bounded*.

The proof is, of course, given by induction. For $\nu = 0$ and $\nu = 1$ (13) is, obviously, true. Let it be true for some value ν, we then have

$$(\Sigma \cdot |\boldsymbol{A}|^{\nu+1}|q) = \sum_{i=0}^{\infty}{}' \sum_{\alpha=0}^{\infty}{}' (i||\boldsymbol{A}|^{\nu}|\alpha)(\alpha||\boldsymbol{A}||q) =$$

$$\sum_{\alpha=0}^{\infty}{}' \left(\sum_{i=0}^{\infty}{}' (i||\boldsymbol{A}|^{\nu}|\alpha) \right) (\alpha||\boldsymbol{A}||q) \leq M^{\nu} M = M^{\nu+1} \qquad (14)$$

for all $q = 0, 1, 2, \ldots$

which proves (13)[1]). (14) shows that all the column sums of $|\boldsymbol{A}|^{\nu+1}$ are convergent and bounded with the bound $M^{\nu+1}$ and thus, eo ipso, that $|\boldsymbol{A}|^{\nu+1}$ exists and is a bounded matrix with the bound $M^{\nu+1}$, q. e. d.

Theorem 1.

If for an arbitrary matrix $\boldsymbol{A}(t)$ $\boldsymbol{K}$ given in (5) *is bounded, then $\boldsymbol{A}(t)$, being of course also bounded, is absolutely exponentiable in each interval (x_0, x). The exponential, being a bounded matrix with the bound* $\exp[M|x-x_0|]$, *is, furthermore, a continuous matrix function*[2]).

In fact, applying lemma 1 to $\boldsymbol{K}$ we have

$$|(\Sigma \cdot \mathbf{exp}[\boldsymbol{A}(t)(t-x_0)]|q)| \leq \sum_{i=0}^{\infty}{}' \sum_{\nu=0}^{\infty}{}' (i|\boldsymbol{K}^{\nu}|q) \frac{|x-x_0|^{\nu}}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}{}' \left(\sum_{i=0}^{\infty}{}' (i|\boldsymbol{K}^{\nu}|q) \right) \frac{|x-x_0|^{\nu}}{\nu!} \leq \sum_{\nu=0}^{\infty}{}' M^{\nu} \frac{|x-x_0|^{\nu}}{\nu!} = \exp[M|x-x_0|] < \infty \qquad (15)$$

for all $q = 0, 1, 2, \ldots$
" " intervals (x_0, x)
" " t in $x_0 \leq t$;

(15) shows that the column sums of $\mathbf{exp}[\boldsymbol{A}(t)(t-x_0)]$ are absolutely convergent and uniformly bounded. Consequently, $\mathbf{exp}[\boldsymbol{A}(t)(t-x_0)]$ exists and is again a bounded matrix. Furthermore, (15) shows that the convergence is uniform in t. The exponential is, consequently, continuous, q. e. d.

---

[1]) The inversion of the order of summation is legitimate in this instance due to the theorem that the order of summation of non-negative double series is irrelevant (cf. e. g. Titchmarsh (1932) p. 29).

[2]) We note that theorem 1 shows that our theory embraces Feller's theory (Feller (1937)) as a special case, as he assumes the probability operator to be a bounded matrix, cf. (1.2.17) and (2.1.14).

§ 7.2. When we now consider non-bounded matrices, which are more important in the practical applications than bounded matrices, cf. chap. 4, matters will, of course, become much more complicated as shown by example 1 § 7.1. We shall here discuss only three types which will, as mentioned, cover most matrices met with in practice.

We first give some definitions[1]).

A matrix we shall call **semi-finite**[2]) when in each row and each column there is only a *finite* number of non-vanishing elements, but not necessarily the *same* number for different rows and columns. If $\boldsymbol{A}$ is a function of a variable x, the number of non-vanishing elements in each row and each column shall be the same for all values of x. If $\boldsymbol{A}$ is only semi-finite in the rows (columns), we shall call $\boldsymbol{A}$ **row (column) semi-finite** and with a generic term **half semi-finite,** writing

$$(p|\boldsymbol{A}|q) = A(p, q)\Delta(0 \leq q \leq {}_p n(\boldsymbol{A})) \tag{1}$$

when $\boldsymbol{A}$ is *row semi-finite* and

$$(p|\boldsymbol{A}|q) = A(p, q)\Delta(0 \leq p \leq n(\boldsymbol{A})_q) \tag{2}$$

when $\boldsymbol{A}$ is *column semi-finite*. Here $A(p, q)$ denotes some function of p and q, and the Δ-symbol is given by (1. 2. 10), i. e.

$$\Delta(0 \leq \alpha \leq a) = \begin{cases} 1 \text{ for } 0 \leq \alpha \leq a \\ 0 \text{ ,, all } \alpha < 0 \text{ and } \alpha > a \,. \end{cases} \tag{3}$$

Finally, ${}_p n(\boldsymbol{A})$ denotes the maximum *column* index in the p'th *row* and $n(\boldsymbol{A})_q$ the maximum *row* index in the q'th *column* of the non-vanishing elements of $\boldsymbol{A}$, i. e.

$$(p|\boldsymbol{A}|j) \begin{cases} \neq 0 \text{ for } j = {}_p n(\boldsymbol{A}) \\ = 0 \text{ ,, all } j > {}_p n(\boldsymbol{A}) \end{cases} \quad \text{for each fixed } p = 0, 1, 2, \ldots \tag{4}$$

$$(i|\boldsymbol{A}|q) \begin{cases} \neq 0 \text{ for } i = n(\boldsymbol{A})_q \\ = 0 \text{ ,, all } i > n(\boldsymbol{A})_q \end{cases} \quad \text{,, ,, ,, } q = 0, 1, 2, \ldots \,. \tag{5}$$

Lemma 1.

For two row (column) semi-finite matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ the product $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C}$ defined in the usual way, (7.1.4), *always exists and is again a row (column) semi-finite matrix.*

---

[1]) In § 7.6 we shall, as mentioned, summarize all our definitions and theorems.

[2]) Toeplitz (1909) has called such matrices *finite*, but we prefer to reserve this term for another purpose, cf. p. 199.

Firstly, all the sums entering into (7. 1. 4) are, namely, *finite* sums[1]) and the product, consequently, exists. Secondly, we have, inserting (1) into (7. 1. 4) in case A and B are both *row* semi-finite,

$$(p|C|q) = \sum_{\alpha=0}^{(\infty)}{}' A(p, \alpha)\Delta(0 \leqq \alpha \leqq {}_p n(A))B(\alpha, q)\Delta(0 \leqq q \leqq {}_\alpha n(B)) = C(p, q)\Delta(0 \leqq q \leqq {}_p n(C)) \quad (6)$$

in which

$${}_p n(C) \leqq \max \{{}_\alpha n(B)\} \text{ for } \alpha = 0, 1, 2, \ldots, {}_p n(A). \quad (7)$$

(6) shows that the product C is again a row semi-finite matrix, q. e. d. Inserting (2) into (7.1.4) in case A and B are both *column* semi-finite, we have in the same way

$$(p|C|q) = \sum_{\alpha=0}^{(\infty)}{}' A(p, \alpha)\Delta(0 \leqq p \leqq n(A)_\alpha)B(\alpha, q)\Delta(0 \leqq \alpha \leqq n(B)_q) = C(p, q)\Delta(0 \leqq p \leqq n(C)_q) \quad (8)$$

in which

$$n(C)_q \leqq \max \{n(A)_\alpha\} \quad \text{for} \quad \alpha = 0, 1, 2, \ldots, n(B)_q. \quad (9)$$

(8) shows that the product C is again a column semi-finite matrix, q. e. d.

For the product of half semi-finite matrices the associative and the distributive (but as usual *not* the commutative) rule is, of course, valid, all the sums entering being finite sums[2]).

As a consequence of lemma 1 we have

Corollary 1.

For an arbitrary half semi-finite matrix all the powers

$$A^\nu \quad \nu = 0, 1, 2, \ldots \quad (10)$$

exist and are again half semi-finite matrices.

Semi-finiteness is, however, in itself, in spite of corollary 1, *not a sufficient condition to ensure* A *to be exponentiable* as shown later in example 1 § 7.3. In one important case it is, however, a sufficient

[1]) When in the following we are not especially interested in keeping account of the number of terms in a sum, we shall express this fact by simply writing $\sum^{(\infty)}$.

[2]) We note that for *arbitrary, infinite* matrices the distributive rule is certainly also valid, but the associative rule generally *not*. A *sufficient* condition for the associative rule is, however, that the product $|A| \cdot |B| \cdot |C|$ exists.

condition. By a **finite matrix** of order N we shall understand a semi-finite matrix for which

$$_p n(A) \leq N \quad \text{for all } p = 0, 1, 2, \ldots \tag{11}$$

and

$$n(A)_q \leq N \quad \text{for all } q = 0, 1, 2, \ldots \tag{12}$$

If only one of the relations (11) and (12) are fulfilled, we shall call A **half finite** of order N[1]. As a consequence of corollary 1 we have

Corollary 2.

For an arbitrary half finite matrix of order N all the powers A^ν, $\nu = 0, 1, 2, \ldots$, exist and are again all half finite matrices of order N if only we define A^0 as the finite unit matrix of order N.

The first part of this corollary follows from corollary 1 and the second part follows from (7) and (9).

Lemma 2.

In case of (11) let $_pM$ denote the maximum of the N first and the p'th numerical row sums, i. e.

$$_pM = \max \left\{(i||A| \cdot \Sigma)\right\} \quad \text{for } i = 0, 1, 2, \ldots, N, p. \tag{13}$$

We then have

$$(p||A|^\nu \cdot \Sigma) \leq {_pM^\nu} \qquad \begin{matrix} \text{for all } p = 0, 1, 2, \ldots \\ \text{,, ,, } \nu = 0, 1, 2, \ldots \end{matrix} \tag{14}$$

The proof is given by a modification of the proof for lemma 1 § 7.1. For $\nu = 0$ and $\nu = 1$ (14) is, obviously, true. Let it be valid for some value ν, we then have due to (11) and corollary 2

$$(p||A|^{\nu+1} \cdot \Sigma) = \sum_{j=0}^{N}{}' \sum_{\alpha=0}^{N}{}' (p||A|^\nu|\alpha)(\alpha||A||j) =$$
$$\sum_{\alpha=0}^{N}{}' (p||A|^\nu|\alpha) \sum_{j=0}^{N}{}' (\alpha||A||j) \leq {_pM^\nu}\, {_pM} = {_pM^{\nu+1}}. \tag{15}$$

Consequently (14) is true for all values of ν, q. e. d.

Lemma 3.

In case of (12) let M_q denote the maximum of the N first and the q'th numerical column sums, i. e.

[1]) We note that in case of (12) it follows from (2.1.11) that $p(n, f) \equiv 0$ for all $n > N$ and, consequently, that (11) must in this case also be fulfilled if A is of the form of a probability operator.

$$M_q = \max\left\{(\Sigma \cdot |A||j)\right\} \quad \text{for } j = 0, 1, 2, \ldots, N, q. \tag{16}$$

We then have

$$(\Sigma \cdot |A|^{\nu}|q) \leqq M_q^{\nu} \qquad \text{for all } q = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \nu = 0, 1, 2, \ldots . \tag{17}$$

For instead of (15) we now have

$$(\Sigma \cdot |A|^{\nu+1}|q) = \sum_{i=0}^{N}{}' \sum_{\alpha=0}^{N}{}' (i||A|^{\nu}|\alpha)(\alpha||A||q) =$$

$$\sum_{\alpha=0}^{N}{}' \left(\sum_{i=0}^{N}{}' (i||A|^{\nu}|\alpha) \right) (\alpha||A||q) \leqq M_q^{\nu} M_q = M_q^{\nu+1}, \tag{18}$$

q. e. d.

Theorem 1.

An arbitrary half finite matrix $A(t)$ of order N is absolutely exponentiable in each interval (x_0, x) and the exponential is again half finite of order N. Furthermore, the exponential is a continuous matrix function and has absolutely convergent row, respectively column, sums.

Firstly, it follows from corollary 2 that if the exponential exist, it is a half finite matrix of order N. Secondly, we have that K given in (7.1.5) is again half finite and we may, therefore, apply lemma 2 and 3 to K. In case of (11) it then follows from (14) that

$$|(p| \exp[A(t)(t-x_0)] \cdot \Sigma)| \leqq \sum_{j=0}^{N}{}' \sum_{\nu=0}^{\infty}{}' (p|K^{\nu}|j) \frac{|x-x_0|^{\nu}}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}{}' \left(\sum_{j=0}^{N}{}' (p|K^{\nu}|j) \right) \frac{|x-x_0|^{\nu}}{\nu!} \leqq \sum_{\nu=0}^{\infty}{}' {}_pM^{\nu} \frac{|x-x_0|^{\nu}}{\nu!} =$$

$$\exp[{}_pM|x-x_0|] < \infty \qquad \text{for all } p = 0, 1, 2, \ldots \quad ,, \quad ,, \text{ intervals } (x_0, x) \quad ,, \quad ,, \; t \text{ in } x_0 \leqq t \leqq x. \tag{19}$$

(19) shows that the row sums of $\exp[A(t)(t-x_0)]$ are absolutely convergent and, eo ipso, that the exponential exists for all intervals. Furthermore, (19) shows that the convergence is uniform in t. The exponential is, consequently, continuous, q. e. d.[1])

[1]) We note that this proof is in our case unnecessary. From (7.1.8) and (11) it follows, namely, that A is a *bounded* matrix. In case of (12) we have observed that (11) follows. In both cases A is, therefore, bounded and theorem 1 is thus contained in theorem 1 § 7.1.

In case of (12) it follows in the same way from (17) that

$$|(\Sigma \cdot \exp [A(t)(t-x_0)]|q)| \leq \sum_{i=0}^{N} \sum_{\nu=0}^{\infty} (i|K^\nu|q) \frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty} \left(\sum_{i=0}^{N} (i|K^\nu|q) \right) \frac{|x-x_0|^\nu}{\nu!} \leq \sum_{\nu=0}^{\infty} M_q^\nu \frac{|x-x_0|^\nu}{\nu!} =$$

$$\exp [M_q|x-x_0|] < \infty \qquad \begin{array}{l} \text{for all } q = 0, 1, 2, \ldots \\ \text{,, ,, intervals } (x_0, x) \\ \text{,, ,, } t \text{ in } x_0 \leq t \leq x . \end{array} \tag{20}$$

(20) shows that the column sums of **exp** $[A(t)(t-x_0)]$ are absolutely convergent and, eo ipso, that the exponential exists for all intervals. Furthermore, (20) shows that the convergence is uniform in t. The exponential is, consequently, continuous, q. e. d.

§ 7.3. We shall call a matrix **semi-diagonal** if it is semi-finite and an integer l, furthermore, exist so that

$$_pn(A) \leq p+l \qquad \text{for all } p = 0, 1, 2, \ldots \tag{1}$$

and

$$n(A)_q \leq q+l \qquad \text{,, ,, } q = 0, 1, 2, \ldots \quad (l = 0, 1, 2, \ldots) . \tag{2}$$

If only (1) *or* (2) is fulfilled, we shall call A **row** or **column semi-diagonal** and with a generic term **half semi-diagonal.** (7.2.1) and (7.2.2) is thus reduced to

$$(p|A|q) = A(p, q)\Delta(0 \leq q \leq p+l) \tag{3}$$

when A is *row semi-diagonal* and

$$(p|A|q) = A(p, q)\Delta(0 \leq p \leq q+l) \tag{4}$$

when A is *column semi-diagonal.*

Lemma 1.

For an arbitrary half semi-diagonal matrix A all the powers A^ν, $\nu = 0, 1, 2, \ldots$, exist and are again half semi-diagonal matrices given by

$$(p|A^\nu|q) = A^{(\nu)}(p, q)\Delta(0 \leq q \leq p+\nu l) \tag{5}$$

when A is row semi-diagonal and

$$(p|A^\nu|q) = A^{(\nu)}(p, q)\Delta(0 \leq p \leq q+\nu l) \tag{6}$$

when A is column semi-diagonal. Thus we have

$$_p n(A^\nu) \leqq p+\nu l \qquad \text{for all } p = 0, 1, 2, \ldots \tag{7}$$

and

$$n(A^\nu)_q \leqq q+\nu l \qquad \text{,, ,, } q = 0, 1, 2, \ldots \tag{8}$$

respectively.

The first part of the lemma follows immediately from corollary 1 § 7.2. The other part follows from (7.2.7) and (7.2.9) by induction. In fact (7) is true for $\nu = 0$ and $\nu = 1$. Let it be true for some value ν, we then have due to (7.2.7) and (1)

$$\begin{aligned} _p n(A^{\nu+1}) &\leqq \max \{_\alpha n(A)\} \quad \text{for } \alpha = 0, 1, 2, \ldots, {_p n(A^\nu)} \\ &\leqq \max \{\alpha+l\} \qquad \text{for } \alpha = 0, 1, 2, \ldots, p+\nu l \\ &= p+\nu l+l = p+(\nu+1)l, \end{aligned} \tag{9}$$

q. e. d. In the same way we have that (8) is true for $\nu = 0$ and $\nu = 1$ and furthermore, using (7.2.9) and (2),

$$\begin{aligned} n(A^{\nu+1})_q &\leqq \max \{n(A)_\alpha\} \quad \text{for } \alpha = 0, 1, 2, \ldots, n(A^\nu)_q \\ &\leqq \max \{\alpha+l\} \qquad \text{for } \alpha = 0, 1, 2, \ldots, q+\nu l \\ &= q+\nu l+l = q+(\nu+1)l, \end{aligned} \tag{10}$$

q. e. d. We note that (7) and (8) respectively, of course, also hold if A is a function of a variable x and we replace A^ν by the product $A(x_1)\cdots A(x_\nu)$ in which $x_1, \ldots, x_\nu$ denote different values of x.

If, especially, A is half semi-diagonal with $l = 0$, i. e.

$$(p|A|q) = A(p, q)\Delta(0 \leqq q \leqq p) \tag{11}$$

and

$$(p|A|q) = A(p, q)\Delta(0 \leqq p \leqq q) \tag{12}$$

respectively, or

$$_p n(A) \leqq p \qquad \text{for all } p = 0, 1, 2, \ldots \tag{13}$$

and

$$n(A)_q \leqq q \qquad \text{,, ,, } q = 0, 1, 2, \ldots \tag{14}$$

respectively, we shall call A a **(row or column) half matrix**[1]).

Corollary 1.

For an arbitrary half matrix all the powers A^ν, $\nu = 0, 1, 2, \ldots$, exist and are again half matrices.

This corollary follows at once from lemma 1 putting $l = 0$ in (5) and (6) respectively.

---

[1]) We note that we only require *one* of the relations (13) and (14) to be fulfilled. If they were both satisfied, A would, namely, be a *diagonal* matrix which case is a trivial one.

Even the property of A being semi-diagonal is not in itself a sufficient condition for A to be exponentiable as shown by the following example.

Example 1.

Let A be a semi-diagonal matrix given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \dots \\ 1 & 0 & 1 & 0 \dots \\ 0 & 1 & 0 & 1 \dots \\ 0 & 0 & 1 & 0 \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} f(0) & 0 & 0 & 0 \dots \\ 0 & f(1) & 0 & 0 \dots \\ 0 & 0 & f(2) & 0 \dots \\ 0 & 0 & 0 & f(3) \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{15}$$

in which $f(n)$ is an arbitrary function of the integer $n = 0, 1, 2, \dots$. We can thus write (15) in the form

$$(p|A|q) = (\delta_{p,q-1} + \delta_{p,q+1})f(q) = (\delta_{p+1,q} + \delta_{p-1,q})f(q) \tag{16}$$

in which as usual

$$\delta_{p,q} = \begin{cases} 1 & \text{for} \quad p = q \\ 0 & \;\;,, \quad\; p \neq q \,. \end{cases} \tag{17}$$

A given in (15) has, of course, absolutely convergent column sums given by

$$(\Sigma \cdot A|q) = 2f(q)\,. \tag{18}$$

From (15) it is easily seen that A^ν will have only non-vanishing elements in the diagonals lying $\pm 1, \pm 3, \dots, \pm \nu$ steps from the main diagonal if ν is odd and $0, \pm 2, \pm 4, \dots, \pm \nu$ if ν is even. It is unnecessary for our purpose (and would be rather cumbersome) to evaluate all the elements of A^ν. It is sufficient to evaluate the two extreme elements in each column and each row of A^ν. In fact, we easily find by induction

$$(p|A^\nu|q) = \delta_{p+\nu,q} \prod_{\beta=0}^{\nu-1} f(q-\beta) + \dots + \delta_{p-\nu,q} \prod_{\beta=0}^{\nu-1} f(q+\beta) \tag{19}$$

for all $p, q = 0, 1, 2, \dots$
$\quad ,, \quad ,, \quad \nu = 1, 2, 3, \dots$.

(19) is, obviously, true for $\nu = 1$. From (16) it then follows that

$$(p|A^{\nu+1}|q) = \sum_{\alpha=0}^{(\infty)} (p|A|\alpha)(\alpha|A^\nu|q) =$$

$$\sum_{\alpha=0}^{(\infty)} \left(\delta_{p+1,\alpha} f(\alpha) + \delta_{p-1,\alpha} f(\alpha)\right)\left(\delta_{\alpha,q-\nu} \prod_{\beta=0}^{\nu-1} f(q-\beta) + \dots + \delta_{\alpha,q+\nu} \prod_{\beta=0}^{\nu-1} f(q+\beta)\right) =$$

$$\delta_{p+1,q-\nu} f(q-\nu) \prod_{\beta=0}^{\nu-1} f(q-\beta) + \dots + \delta_{p-1,q+\nu} f(q+\nu) \prod_{\beta=0}^{\nu-1} f(q+\beta) =$$

$$\delta_{p+(\nu+1),q} \prod_{\beta=0}^{(\nu+1)-1} f(q-\beta) + \dots + \delta_{p-(\nu+1),q} \prod_{\beta=0}^{(\nu+1)-1} f(q+\beta)\,. \tag{20}$$

as all other terms will not contribute to the two extreme terms, q. e. d. From (19) we now find

$$(p|A^{2\nu}|q) = \sum_{\alpha=0}^{(\infty)}{}' (p|A^{\nu}|\alpha)(\alpha|A^{\nu}|q) =$$

$$\sum_{\alpha=0}^{(\infty)} \left(\delta_{p+\nu,\alpha} \prod_{\beta=0}^{\nu-1} f(\alpha-\beta) + \dots\right)\left(\dots + \delta_{\alpha,q+\nu} \prod_{\beta=0}^{\nu-1} f(q+\beta)\right) =$$

$$\dots + \delta_{p,q}\left(\prod_{\beta=1}^{\nu} f(q+\beta)\right)\left(\prod_{\beta=0}^{\nu-1} f(q+\beta)\right) + \dots \quad . \tag{21}$$

As we have here omitted several *positive* terms, even in the coefficient of $\delta_{p,q}$, it follows from (21) that

$$(q|A^{2\nu}|q)\frac{(x-x_0)^{2\nu}}{(2\nu)!} > \left(\prod_{\beta=1}^{\nu} f(q+\beta)\right)\left(\prod_{\beta=0}^{\nu-1} f(q+\beta)\right)\frac{(x-x_0)^{2\nu}}{(2\nu)!} \tag{22}$$

for all $q = 0, 1, 2, \dots$.

If e. g. we now put

$$f(n) = n^{1+r} \quad (r > 0), \tag{23}$$

we find by means of the Stirling formula for the factorial

$$(q|A^{2\nu}|q)\frac{(x-x_0)^{2\nu}}{(2\nu)!} > ((\nu+1)!\,\nu!)^{1+r}\frac{(x-x_0)^{2\nu}}{(2\nu)!} \sim \text{const.}\ \nu^{\frac{3}{2}+2r}\left(\frac{\nu(x-x_0)^{\frac{1}{r}}}{2^{\frac{1}{r}}e}\right)^{2r\nu} \underset{\nu\to\infty}{\longrightarrow} \infty$$

for all $q \geq 1$
" " $r > 0$. (24)

(24) shows that $\mathbf{exp}\,[A(x-x_0)]$ does *not* exist on the assumption (23), as even the individual term of the series diverges, q. e. d.

Lemma 2.

Let A be a row semi-diagonal matrix, characterized by (1). *Let ${}_sM$ denote the maximum of the s first absolute row sums of A, i. e.*

$${}_sM = \max\{(\alpha||A|\cdot\Sigma)\} \quad \text{for } \alpha = 0, 1, 2, \dots, s\,. \tag{25}$$

We then have

$$(p||A|^{\nu}\cdot\Sigma) \leq \prod_{\beta=0}^{\nu-1} {}_{p+\beta}M \qquad \text{for all } p = 0, 1, 2, \dots \quad \text{" " } \nu = 1, 2, 3, \dots \,. \tag{26}$$

The proof is, of course, performed by a modification of (7.1.14) and (7.2.15). For $\nu = 1$ (26) is, obviously, true. Let it be true for some value of ν, we then have, using (7),

$$(p||A|^{\nu+1}\cdot\Sigma) = \sum_{j=0}^{(\infty)}\sum_{\alpha=0}^{p+\nu l}(p||A|^{\nu}|\alpha)(\alpha||A||j) =$$

$$\sum_{\alpha=0}^{p+\nu l}(p||A|^{\nu}|\alpha)\sum_{j=0}^{(\infty)}(\alpha||A||j) \leq \left(\prod_{\beta=0}^{\nu-1}{}_{p+\beta l}M\right){}_{p+\nu l}M = \prod_{\beta=0}^{(\nu+1)-1}{}_{p+\beta l}M\,. \quad (27)$$

(26) is, consequently, true for all values of ν, q. e. d.

Corollary 2.

For a row half matrix, characterized by (13), *we have*

$$(p||A|^{\nu}\cdot\Sigma) \leq {}_{p}M^{\nu} \qquad \text{for all } p = 0, 1, 2, \ldots \quad \nu = 0, 1, 2, \ldots \quad (28)$$

For $\nu = 0$ (28) is, obviously, true and for $\nu > 0$ it follows immediately from (26) putting $l = 0$.

Lemma 3.

Let A be a column semi-diagonal matrix, characterized by (2). *Let M_s denote the maximum of the s first absolute column sums of A, i. e.*

$$M_s = \max\{(\Sigma\cdot|A||\alpha)\} \quad \text{for } \alpha = 0, 1, 2, \ldots, s\,. \quad (29)$$

We then have

$$(\Sigma\cdot|A|^{\nu}|q) \leq \prod_{\beta=0}^{\nu-1} M_{q+\beta l} \qquad \text{for all } q = 0, 1, 2, \ldots \quad \nu = 1, 2, 3, \ldots \quad (30)$$

For $\nu = 1$ (30) is, obviously, true and instead of (27) we now have, using (8),

$$(\Sigma\cdot|A|^{\nu+1}|q) = \sum_{i=0}^{(\infty)}\sum_{\alpha=0}^{q+\nu l}(i||A||\alpha)(\alpha||A|^{\nu}|q) =$$

$$\sum_{\alpha=0}^{q+\nu l}\left(\sum_{i=0}^{(\infty)}(i||A||\alpha)\right)(\alpha||A|^{\nu}|q) \leq M_{q+\nu l}\prod_{\beta=0}^{\nu-1}M_{q+\beta l} = \prod_{\beta=0}^{(\nu+1)-1}M_{q+\beta l}\,. \quad (31)$$

(30) is, consequently, true for all values of ν, q. e. d.

Corollary 3.

For a column half matrix, characterized by (14), *we have*

$$(\Sigma\cdot|A|^{\nu}|q) \leq M_q^{\nu} \qquad \text{for all } q = 0, 1, 2, \ldots \quad \nu = 0, 1, 2, \ldots \quad (32)$$

For $\nu = 0$ (32) is, obviously, true and for $\nu > 0$ it follows immediately from (30) putting $l = 0$.

We now see from corollaries 2 and 3 that in one important case the property of A being half semi-diagonal is sufficient to ensure that A is exponentiable, namely if A is a half matrix.

Theorem 1.

An arbitrary half matrix $A(t)$ is absolutely exponentiable in each interval (x_0, x) and the exponential is again a half matrix. Furthermore, it is a continuous matrix function and has absolutely convergent row, respectively column, sums.

Firstly, it follows from corollary 1 that if the exponential exist, it is a half matrix. Next, we have, in case A is a *row* half matrix, that also K given in (7.1.5) is a row half matrix. Applying (28) to K we now have

$$\left|(p|\exp\left[A(t)(t-x_0)\right]\cdot\Sigma)\right| \leq \sum_{j=0}^{p}{}' \sum_{\nu=0}^{\infty}{}' (p|K^\nu|j)\frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}{}'\left(\sum_{j=0}^{p}{}'(p|K^\nu|j)\right)\frac{|x-x_0|^\nu}{\nu!} \leq \sum_{\nu=0}^{\infty}{}' {}_pM^\nu\frac{|x-x_0|^\nu}{\nu!} = \exp\left[{}_pM|x-x_0|\right] < \infty$$

$$\begin{array}{ll} \text{for all } p = 0, 1, 2, \ldots & \\ \text{,, ,, intervals } (x_0, x) & \quad (33) \\ \text{,, ,, } t \text{ in } x_0 \leqq t \leqq x\,. & \end{array}$$

In the same way we have from (32) in case A is a *column* half matrix

$$\left|(\Sigma\cdot\exp\left[A(t)(t-x_0)\right]|q)\right| \leq \sum_{i=0}^{q}{}' \sum_{\nu=0}^{\infty}{}' (i|K^\nu|q)\frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}{}'\left(\sum_{i=0}^{q}{}'(i|K^\nu|q)\right)\frac{|x-x_0|^\nu}{\nu!} \leq \sum_{\nu=0}^{\infty}{}' M_q^\nu\frac{|x-x_0|^\nu}{\nu!} = \exp\left[M_q|x-x_0|\right] < \infty$$

$$\begin{array}{ll} \text{for all } q = 0, 1, 2, \ldots & \\ \text{,, ,, intervals } (x_0, x) & \quad (34) \\ \text{,, ,, } t \text{ in } x_0 \leqq t \leqq x\,. & \end{array}$$

(33) and (34) show that all the row and column sums, respectively, of $\exp\left[A(t)(t-x_0)\right]$ are absolutely convergent and, eo ipso, that the exponential exists. Furthermore, they show that the convergence is uniform in t and the exponential is, consequently, continuous, q.e.d.

§ 7.4. We shall finally discuss the last case, which is the most important one for our applications of the theory (cf. chap. 4), namely that of A being *column semi-diagonal and the absolute column sums increasing at most proportionately with* q, i.e.

$$n(A)_q \leqq q+l \tag{1}$$

and

$$(\Sigma \cdot |A(t)| |q) \leqq f(t) \cdot q \quad \text{for all } q \geqq 1 . \tag{2}$$

Theorem 1.

An arbitrary column semi-diagonal matrix $A(t)$ satisfying (2) *is absolutely exponentiable in each interval* (x_0, x) *for which*

$$|x - x_0| < \frac{1}{lC} \tag{3}$$

with

$$C = \max_{x_0 \leq t < x} f(t) . \tag{4}$$

Furthermore, the exponential is continuous and has absolutely convergent column sums.

From (2) and (4) we have, applying lemma 3 § 7.3 to K given in (7.1.5),

$$\left|(\Sigma \cdot \exp\left[A(t)(t-x_0)\right]|q)\right| \leqq \sum_{i=0}^{\infty} \sum_{\nu=0}^{\infty} (i|K^\nu|q) \frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}\left(\sum_{i=0}^{(\infty)} (i|K^\nu|q)\right) \frac{|x-x_0|^\nu}{\nu!} \leqq \sum_{\nu=0}^{\infty}\left(\prod_{\beta=0}^{\nu-1} M_{q+\beta l}\right) \frac{|x-x_0|^\nu}{\nu!} \leqq$$

$$\sum_{\nu=0}^{\infty} M_q \left(\prod_{\beta=1}^{\nu-1} (q+\beta l)\right) C^{\nu-1} \frac{|x-x_0|^\nu}{\nu!} = \sum_{\nu=0}^{\infty} u_\nu \tag{5}$$

for all $q = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, t in $x_0 \leqq t \leqq x$.

(5) is, however, convergent for all intervals satisfying (3) because

$$\frac{u_{\nu+1}}{u_\nu} = \frac{q+\nu l}{\nu+1} C|x-x_0| \underset{\nu\to\infty}{\longrightarrow} lC|x-x_0| < 1 . \tag{6}$$

Furthermore, it is seen that the series (5) is uniformly convergent in $x_0 \leqq t \leqq x$. The exponential is, consequently, continuous, q. e. d.[1])

It might be hoped that the condition (3) would be too narrow, i. e. that the majorization in (5) would be too rough. *They can, however, generally not be essentially improved* as shown by the following example.

---

[1]) We note that (5) and (6) are also valid for $l = 0$. In this case the assumption (2) is, however, unnecessary as shown in theorem 1 § 7.3.

Example 1.

Let A be the following semi-diagonal matrix

$$A = \begin{Bmatrix} 1 & 0 & 0 & 0\ldots \\ 1 & 1 & 0 & 0\ldots \\ 0 & 1 & 1 & 0\ldots \\ 0 & 0 & 1 & 1\ldots \\ \vdots & \vdots & \vdots & \vdots \end{Bmatrix} \cdot \begin{Bmatrix} 0 & 0 & 0 & 0\ldots \\ 0 & 1 & 0 & 0\ldots \\ 0 & 0 & 2 & 0\ldots \\ 0 & 0 & 0 & 3\ldots \\ \vdots & \vdots & \vdots & \vdots \end{Bmatrix} \tag{7}$$

i. e.

$$(p|A|q) = (\delta_{p,q} + \delta_{p-1,q})q \qquad p, q = 0, 1, 2, \ldots \ . \tag{8}$$

The conditions of theorem 1 are thus fulfilled with $l = 1$ and $f(x) \equiv 2$. $\exp[A(x-x_0)]$ consequently exists for all intervals for which

$$|x - x_0| < \frac{1}{1 \cdot 2} = 0{\cdot}5 \, . \tag{9}$$

In this case the exponential may, however, be evaluated direct as it is *the* solution of the differential equation

$$Y' = A \cdot Y \tag{10}$$

which for $x = x_0$ is equal to 1. Inserting (7) into (10) we have simply

$$\frac{d}{dx}(p|Y|q) = (p-1)(p-1|Y|q) + p(p|Y|q) \qquad p, q = 0, 1, 2, \ldots \ . \tag{11}$$

As shown in § 2.6 (11) may be integrated successively, as A is also a *row half matrix*, and for $q = 1$ the result is, as is easily verified,

$$(p|Y|1) = (p|\exp[A(x-x_0)]|1) = \begin{cases} e^{x-x_0}(e^{x-x_0}-1)^{p-1} & \text{for all } p \geqq 1 \\ 0 & \text{,, } p = 0 \ . \end{cases} \tag{12}$$

From (12) it follows that

$$(\Sigma \cdot \exp[A(x-x_0)]|1) =$$

$$e^{x-x_0} \sum_{i=1}^{\infty} (e^{x-x_0}-1)^{i-1} = \begin{cases} \dfrac{e^{x-x_0}}{2-e^{x-x_0}} & \text{for } x - x_0 < \ln 2 \\ \infty & \text{,, } x - x_0 \geqq \ln 2 \ . \end{cases} \tag{13}$$

(13) shows that the exponential has for $q = 1$ an absolutely convergent column sum for all intervals satisfying $x - x_0 < \ln 2 = 0{\cdot}69$ which is not materially larger than the limit (9) given by the condition (3)[1].

We finally observe that example 1 § 7.3 shows that the condition (2) in theorem 1 cannot be weakened as the column sums in the example, given by (7.3.18) and (7.3.23), increase only with

[1]) We note that in example 1 the exponential itself exists for *all* intervals whether or not the column sums are convergent. The reason for this is, however, that A besides being column semi-diagonal is also a row half matrix.

q as q^{1+r} where $r(>0)$ may be arbitrarily small. *The general theory cannot, consequently, be pushed beyond theorem* 1. It is interesting to note that *this theoretical limit just covers the matrices most important in practice* (cf. chap. 4). On the other hand, we can easily construct special matrices for which neither (1) nor (2) are satisfied and which are, nevertheless, absolutely exponentiable in every interval, the exponential having, furthermore, absolutely convergent column sums as shown by the following example.

Example 2.

Let $\boldsymbol{A}$ be the following non-semi-finite matrix

$$\boldsymbol{A} = \begin{Vmatrix} 1 & 1 & 1 & 1 & \dots \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \dots \\ (\frac{1}{4})^2 & (\frac{1}{4})^2 & (\frac{1}{4})^2 & (\frac{1}{4})^2 & \dots \\ (\frac{1}{4})^3 & (\frac{1}{4})^3 & (\frac{1}{4})^3 & (\frac{1}{4})^3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix} \cdot \begin{Vmatrix} 1 & 0 & 0 & 0 & \dots \\ 0 & 2 & 0 & 0 & \dots \\ 0 & 0 & 2^2 & 0 & \dots \\ 0 & 0 & 0 & 2^3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}, \tag{14}$$

i. e.

$$(p|\boldsymbol{A}|q) = (\tfrac{1}{4})^p 2^q \qquad p, q = 0, 1, 2, \dots . \tag{15}$$

$\boldsymbol{A}$ thus has convergent column sums because

$$(\Sigma \cdot \boldsymbol{A}|q) = \sum_{i=0}^{\infty} (\tfrac{1}{4})^i 2^q = \frac{1}{1-\frac{1}{4}} 2^q = \frac{4}{3} 2^q < \infty \qquad q = 0, 1, 2, \dots \tag{16}$$

which sums increase with q as 2^q, i. e. much more rapidly than in (2). All the powers of $\boldsymbol{A}$ exist and have the elements

$$(p|\boldsymbol{A}^\nu|q) = (\tfrac{1}{4})^p 2^{q+\nu-1} \qquad \nu = 1, 2, 3, \dots . \tag{17}$$

(17) is, namely, true for $\nu = 1$ and from (15) and (17) we have

$$\begin{aligned} (p|\boldsymbol{A}^{\nu+1}|q) &= \sum_{\alpha=0}^{\infty} (p|\boldsymbol{A}^\nu|\alpha)(\alpha|\boldsymbol{A}|q) = \sum_{\alpha=0}^{\infty} (\tfrac{1}{4})^p 2^{\alpha+\nu-1} (\tfrac{1}{4})^\alpha 2^q \\ &= (\tfrac{1}{4})^p 2^{q+\nu-1} \sum_{\alpha=0}^{\infty} (\tfrac{1}{2})^\alpha = (\tfrac{1}{4})^p 2^{q+(\nu+1)-1} , \end{aligned} \tag{18}$$

q. e. d. From (17) follows that $\boldsymbol{A}^\nu$ has convergent column sums

$$(\Sigma \cdot \boldsymbol{A}^\nu|q) = \tfrac{4}{3} 2^{q+\nu-1} . \tag{19}$$

It is now easily seen that $\boldsymbol{A}$ is absolutely exponentiable in each interval, the exponential having convergent column sums. In fact we have from (17)

$$\begin{aligned} (p|\exp[\boldsymbol{A}(x-x_0)]|q) &= \delta_{p,q} + \sum_{\nu=1}^{\infty} (\tfrac{1}{4})^p 2^{q+\nu-1} \frac{(x-x_0)^\nu}{\nu!} \\ &= \delta_{p,q} + (\tfrac{1}{4})^p 2^{q-1} (\exp[2(x-x_0)] - 1) < \infty \end{aligned} \tag{20}$$

and

$$(\Sigma \cdot \exp[\boldsymbol{A}(x-x_0)]|q) = 1 + \tfrac{4}{3} 2^{q-1} (\exp[2(x-x_0)] - 1) < \infty \tag{21}$$

q. e. d.

§ 7.5. We have so far considered only one-dimensional matrices. We shall now generalize the results obtained for one-dimensional to multi-dimensional matrices. By an **m-dimensional matrix** we understand an (infinite) matrix with m independent row indices, $p_1, \ldots, p_m = (p)$, a point in an m-dimensional p-space, and m independent column indices, $q_1, \ldots, q_m = (q)$, a point in an m-dimensional q-space, i. e. a matrix with elements of the form

$$(p_1, \ldots, p_m|\boldsymbol{A}|q_1, \ldots, q_m) = ((p)|\boldsymbol{A}|(q)) \quad \left.\begin{matrix} p_1, \ldots, p_m \\ q_1, \ldots, q_m \end{matrix}\right\} = 0, 1, 2, \ldots \quad (1)$$

For such matrices all the usual algebraic matrix operations can be defined. In particular the product

$$\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{C} \quad (2)$$

is defined by

$$((p)|\boldsymbol{C}|(q)) = \sum_{\alpha_1=0}^{\infty} \ldots \sum_{\alpha_m=0}^{\infty} ((p)|\boldsymbol{A}|(\alpha))\,((\alpha)|\boldsymbol{B}|(q)) \,. \quad (3)$$

In the following we shall tacitly assume all such summations to be absolutely convergent, i. e. that the order of summation is irrelevant.

The theory of multi-dimensional matrices is, however, contained in the preceding theory of one-dimensional matrices. As is well-known, *the conception of dimension is, namely, as already discussed in chap.* 3 *an artificial conception from the point of view of the theory of manifolds* as shown by the fundamental theorem

$$(\text{enumerable infinite})^m = \text{enumerable infinite} \,. \quad (4)$$

Due to (4) *we can simply re-number both the row and the column indices* 0, 1, 2, ... *and thus transform the m-dimensional matrix into a one-dimensional one.* This re-numbering must, however, be the same for the p- and the q-space (cf. [1]). Because of the fundamental property of absolutely convergent manifold series that the order of summation is irrelevant, *this re-numbering will not influence the result of the summation in* (3)[1]. On the other hand we may, of course, by the inverse process transform a one-dimensional matrix into an m-dimensional one and thus all the examples in §§ 7.1–7.4 may be transformed into corresponding ones for multi-dimensional matrices.

[1]) We note that this fact is, of course, true only if the re-numbering of the p- and the q-space is *identical*. Otherwise, the factors in (3) will not have the same partners after the re-numbering.

We shall now generalize our definitions and theorems from the one- to the multi-dimensional case[1]).

We firstly note that the notions **exponentiable, absolutely exponentiable** and **absolutely exponentiable in an interval** can be applied to m-dimensional matrices direct as the definitions (cf. § 7.1) are independent of the dimension.

By an m-dimensional, **bounded** matrix we understand an m-dimensional matrix for which the numerical column sums are uniformly bounded, i. e.. using (3.1.10).

$$0 \leq \left(\Sigma \cdot |\boldsymbol{A}|(q)\right) \leq M < \infty \quad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \tag{5}$$

in which M is a constant. We can now generalize theorem 1 § 7.1 to the following theorem.

Theorem 1.

If for an arbitrary m-dimensional matrix $\boldsymbol{A}(t)$ $\boldsymbol{K}$ *given in* (7.1.5) *is bounded, then* $\boldsymbol{A}(t)$, *being of course also bounded, is absolutely exponentiable in each interval* (x_0, x). *The exponential, being a bounded matrix with the bound* $\exp\left[M|x-x_0|\right]$, *is, furthermore, a continuous matrix function.*

This theorem follows at once from theorem 1 § 7.1 if we only observe that by any re-numbering transforming $\boldsymbol{A}(t)$ into a one-dimensional matrix the relation (5) will still be satisfied, i. e. the matrix will remain bounded.

By an **m-dimensional semi-finite matrix** we shall understand an m dimensional matrix which for each fixed p-point and each fixed q-point has only a *finite* number of non-vanishing elements. *This fact we shall express briefly by saying that for each fixed p point there is only a finite number of non-vanishing q-points and vice versa.* If $\boldsymbol{A}$ is only semi-finite in the *rows* (*columns*) we shall again call $\boldsymbol{A}$ **row (column) semi-finite** and with a generic term **half semi-finite.** We can now enclose all the non-vanishing q- (p-) points in an m-dimensional interval and, using the Δ-symbol (7.2.3), we can thus write, as a generalization of (7.2.1) and (7.2.2),

$$\left((p)|\boldsymbol{A}|(q)\right) = A\left((p),(q)\right) \prod_{i=1}^{m} \Delta\left(0 \leq q_i < {}_{(p)}n_i(\boldsymbol{A})\right) \tag{6}$$

when $\boldsymbol{A}$ is *row* semi-finite and

$$\left((p)|\boldsymbol{A}|(q)\right) = A\left((p),(q)\right) \prod_{i=1}^{m} \Delta\left(0 \leq p_i < n_i(\boldsymbol{A})_{(q)}\right) \tag{7}$$

[1]) In § 7.6 we shall summarize all our definitions and theorems.

when $\boldsymbol{A}$ is *column* semi-finite. Here $A((p), (q))$ denotes some function of $p_1, \ldots, p_m, q_1, \ldots, q_m$. Furthermore, ${}_{(p)}n_i(\boldsymbol{A})$ denotes the maximum value of the index q_i for fixed $p_1, \ldots, p_m$ and $n_i(\boldsymbol{A})_{(q)}$ the maximum value of the index p_i for fixed $q_1, \ldots, q_m$ of the non-vanishing elements of $\boldsymbol{A}$.

Lemma 1.

For two m-dimensional row (column) semi-finite matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ the product (2) defined in (3) always exists and is again a row (column) semi-finite matrix which can be written in the form (6) and (7) respectively, with

$$ {}_{(p)}n_i(\boldsymbol{C}) \leq \max \left\{ {}_{(\alpha)}n_i(\boldsymbol{B}) \right\} \qquad \text{for} \begin{cases} \alpha_1 = 0, 1, 2, \ldots, {}_{(p)}n_1(\boldsymbol{A}) \\ \cdots\cdots\cdots\cdots \\ \alpha_m = 0, 1, 2, \ldots, {}_{(p)}n_m(\boldsymbol{A}) \end{cases} \tag{8} $$

$$ (i = 1, 2, 3, \ldots, m) $$

and

$$ n_i(\boldsymbol{C})_{(q)} \leq \max \left\{ n_i(\boldsymbol{A})_{(\alpha)} \right\} \qquad \text{for} \begin{cases} \alpha_1 = 0, 1, 2, \ldots, n_1(\boldsymbol{B})_{(q)} \\ \cdots\cdots\cdots\cdots \\ \alpha_m = 0, 1, 2, \ldots, n_m(\boldsymbol{B})_{(q)} \end{cases} \tag{9} $$

$$ (i = 1, 2, 3, \ldots, m) $$

respectively.

The first part of lemma 1 follows immediately from lemma 1 § 7.2 because $\boldsymbol{A}$ will remain half semi-finite by any arbitrary re-numbering transforming $\boldsymbol{A}$ into a one-dimensional matrix. The second part of the lemma, eqs. (8)–(9), follows in analogy with (7.2.7) and (7.2.9) immediately by inserting (6) and (7) into the definition (3):

$$ ((p)|\boldsymbol{C}|(q)) = \sum_{\alpha_1=0}^{(\infty)} \cdots \sum_{\alpha_m=0}^{(\infty)} A((p), (\alpha)) \prod_{i=1}^{m} \Delta(0 \leq \alpha_i \leq {}_{(p)}n_i(\boldsymbol{A})) \cdot $$
$$ B((\alpha), (q)) \prod_{i=1}^{m} \Delta(0 \leq q_i \leq {}_{(\alpha)}n_i(\boldsymbol{B})) \tag{10} $$

in case $\boldsymbol{A}$ and $\boldsymbol{B}$ are both *row* semi-finite. Comparing corresponding Δ-symbols, (8) follows at once. In the same way we obtain in case $\boldsymbol{A}$ and $\boldsymbol{B}$ are both *column* semi-finite

$$ ((p)|\boldsymbol{C}|(q)) = \sum_{\alpha_1=0}^{(\infty)} \cdots \sum_{\alpha_m=0}^{(\infty)} A((p), (\alpha)) \prod_{i=1}^{m} \Delta(0 \leq p_i \leq n_i(\boldsymbol{A})_{(\alpha)}) \cdot $$
$$ B((\alpha), (q)) \prod_{i=1}^{m} \Delta(0 \leq \alpha_i \leq n_i(\boldsymbol{B})_{(q)}). \tag{11} $$

Again comparing corresponding Δ-symbols, (9) follows at once.

As a consequence of lemma 1 we have

Corollary 1.

For an arbitrary m-dimensional half semi-finite matrix $\boldsymbol{A}$ all the powers

$$\boldsymbol{A}^{\nu} \quad \nu = 0, 1, 2, \ldots \qquad (\boldsymbol{A}^{0} = \mathbf{1}) \tag{12}$$

exist and are again half semi-finite matrices.

By an **m-dimensional finite matrix of order N** we shall understand an m-dimensional semi-finite matrix for which

$$_{(p)}n_i(\boldsymbol{A}) \leq N \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \ ,, \ i = 1, 2, 3, \ldots, m \tag{13}$$

and

$$n_i(\boldsymbol{A})_{(q)} \leq N \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \ ,, \ i = 1, 2, 3, \ldots, m . \tag{14}$$

If only one of the conditions (13) and (14) are fulfilled we shall again call $\boldsymbol{A}$ **half finite of order N.**

Corollary 2.

For an arbitrary m-dimensional half finite matrix $\boldsymbol{A}$ all the powers $\boldsymbol{A}^{\nu}$, $\nu = 0, 1, 2, \ldots$, exist and are again all half finite matrices of order N if only we define $\boldsymbol{A}^0$ as the m-dimensional finite unit matrix of order N.

The first part of this corollary follows from corollary 1 and the second part immediately from (8) and (9).

We can now generalize theorem 1 § 7.2 to the following theorem.

Theorem 2.

An arbitrary m-dimensional half finite matrix $\boldsymbol{A}(t)$ of order N is absolutely exponentiable in each interval (x_0, x) and the exponential is again half finite of order N. Furthermore, the exponential is a continuous matrix function and has absolutely convergent row, respectively column, sums.

Firstly, it follows from corollary 2 that if the exponential exist, it is an m-dimensional half finite matrix of order N. The rest of the theorem now follows from theorem 1 § 7.2 because $\boldsymbol{A}(t)$ will remain half finite (although not, of course, necessarily of the *same* order N) by any re-numbering transforming $\boldsymbol{A}(t)$ into a one-dimensional matrix.

An m-dimensional matrix $\boldsymbol{A}$ we shall call **semi-diagonal** if it is semi-finite and, furthermore, an integer l exist so that

$$_{(p)}n_i(\boldsymbol{A}) \leq p_i + l \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{15}$$

and

$$n_i(\boldsymbol{A})_{(q)} \leq q_i + l \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{16}$$

$$(l = 0, 1, 2, \ldots)\ .$$

If only (15) *or* (16) is satisfied, we shall call $\boldsymbol{A}$ **row (column) semi-diagonal** and with a generic term **half semi-diagonal.** (6) and (7) are thus reduced to

$$((p)|\boldsymbol{A}|(q)) = A((p),(q)) \prod_{i=1}^{m} \Delta(0 \leq q_i \leq p_i + l) \tag{17}$$

when $\boldsymbol{A}$ is *row* semi-diagonal and

$$((p)|\boldsymbol{A}|(q)) = A((p),(q)) \prod_{i=1}^{m} \Delta(0 \leq p_i \leq q_i + l) \tag{18}$$

when $\boldsymbol{A}$ is *column* semi-diagonal.

Lemma 2.

For an arbitrary m-dimensional half semi-diagonal matrix $\boldsymbol{A}$ all the powers $\boldsymbol{A}^{\nu}$, $\nu = 0, 1, 2, \ldots$, exist and are again half semi-diagonal matrices given by

$$((p)|\boldsymbol{A}^{\nu}|(q)) = A^{(\nu)}((p),(q)) \prod_{i=1}^{m} \Delta(0 \leq q_i \leq p_i + \nu l) \tag{19}$$

when $\boldsymbol{A}$ is row semi-diagonal and

$$((p)|\boldsymbol{A}^{\nu}|(q)) = A^{(\nu)}((p),(q)) \prod_{i=1}^{m} \Delta(0 \leq p_i \leq q_i + \nu l) \tag{20}$$

when $\boldsymbol{A}$ is column semi-diagonal. Thus, we have

$$_{(p)}n_i(\boldsymbol{A}^{\nu}) \leq p_i + \nu l \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{21}$$

and

$$n_i(\boldsymbol{A}^{\nu})_{(q)} \leq q_i + \nu l \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{22}$$

respectively.

The first part of the lemma follows from corollary 1. The other part follows from (8) and (9) by induction in exactly the same way in which (7.3.7) and (7.3.8) were proved.

Let, especially, $\boldsymbol{A}$ be an m-dimensional *half* semi-diagonal matrix with $l = 0$, i. e.

$$((p)|A|(q)) = A((p),(q)) \prod_{i=1}^{m} \Delta(0 \leqq q_i \leqq p_i) \tag{23}$$

and

$$((p)|A|(q)) = A((p),(q)) \prod_{i=1}^{m} \Delta(0 \leqq p_i \leqq q_i) \tag{24}$$

respectively, or expressed in another way

$$_{(p)}n_i(A) \leqq p_i \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{25}$$

and

$$n_i(A)_{(q)} \leqq q_i \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad i = 1, 2, 3, \ldots, m \tag{26}$$

respectively. In this case we shall call A an **m-dimensional (row or column) half matrix**. (If both relations are satisfied, A will again be reduced to the trivial case of a *diagonal* matrix.)

Corollary 3.

For an arbitrary m-dimensional half matrix A *all the powers* A^ν, $\nu = 0, 1, 2, \ldots$, *exist and are again half matrices.*

This corollary follows at once from lemma 2 putting $l = 0$ in (21) and (22), respectively.

Lemma 3.

Let A *be an m-dimensional row semi-diagonal matrix, characterized by* (15). *Let* $_{(s)}M$ *denote the m-dimensional generalization of* (7.3.25), *i. e.*

$$_{(s)}M = \max \left\{ ((\alpha)||A| \cdot \Sigma) \right\} \quad \text{for} \begin{cases} \alpha_1 = 0, 1, 2, \ldots, s_1 \\ \cdots\cdots\cdots\cdots \\ \alpha_m = 0, 1, 2, \ldots, s_m . \end{cases} \tag{27}$$

We then have

$$((p)||A|^\nu \cdot \Sigma) \leqq \prod_{\beta=0}^{\nu-1} {}_{(p+\beta l)}M \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \nu = 1, 2, 3, \ldots \tag{28}$$

The proof is an immediate generalization of the proof of (7.3.26). For $\nu = 1$ (28) is, obviously, true. Let it be true for some value of ν, we then have, using (21),

$$((p)||A|^{\nu+1} \cdot \Sigma) = \sum_{j_1=0}^{(\infty)} \ldots \sum_{j_m=0}^{(\infty)} \sum_{\alpha_1=0}^{p_1+\nu l} \ldots \sum_{\alpha_m=0}^{p_m+\nu l} ((p)||A|^\nu|(\alpha))((\alpha)||A||(j)) =$$

$$\sum_{\alpha_1=0}^{p_1+\nu l} \ldots \sum_{\alpha_m=0}^{p_m+\nu l} ((p)||A|^\nu|(\alpha)) \sum_{j_1=0}^{(\infty)} \ldots \sum_{j_m=0}^{(\infty)} ((\alpha)||A||(j)) \leqq$$

$$\left(\prod_{\beta=0}^{\nu-1} {}_{(p+\beta l)}M \right) {}_{(p+\nu l)}M = \prod_{\beta=0}^{(\nu+1)-1} {}_{(p+\beta l)}M . \tag{29}$$

(28) is, consequently, true for all values of ν, q. e. d.

Corollary 4.

For an m-dimensional row half matrix $\boldsymbol{A}$, *characterized by* (25), *we have*

$$((p)\|\boldsymbol{A}|^{\nu}\cdot\Sigma) \leq ({}_{(p)}M)^{\nu} \qquad \text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \nu = 0, 1, 2, \ldots \tag{30}$$

For $\nu = 0$ (30) is, obviously, true and for $\nu > 0$ it follows immediately from (28) putting $l = 0$.

Lemma 4.

Let $\boldsymbol{A}$ *be an m-dimensional column semi-diagonal matrix, characterized by* (16). *Let* $M_{(s)}$ *denote the m-dimensional generalization of* (7.3.29), *i. e.*

$$M_{(s)} = \max\left\{(\Sigma\cdot|\boldsymbol{A}|\,|(q))\right\} \quad \text{for} \begin{cases} \alpha_1 = 0, 1, 2, \ldots, s_1 \\ \cdots\cdots\cdots\cdots\cdots \\ \alpha_m = 0, 1, 2, \ldots, s_m \,. \end{cases} \tag{31}$$

We then have

$$(\Sigma\cdot|\boldsymbol{A}|^{\nu}|(q)) \leq \prod_{\beta=0}^{\nu-1} M_{(q+\beta l)} \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \nu = 1, 2, 3, \ldots \tag{32}$$

For $\nu = 1$ (32) is, obviously, true and instead of (29) we now have, using (22), as a generalization of (7.3.31)

$$(\Sigma\cdot|\boldsymbol{A}|^{\nu+1}|(q)) = \sum_{i_1=0}^{(\infty)}\cdots\sum_{i_m=0}^{(\infty)}\sum_{\alpha_1=0}^{q_1+\nu l}\cdots\sum_{\alpha_m=0}^{q_m+\nu l}((i)|\boldsymbol{A}|\,|(\alpha))((\alpha)|\boldsymbol{A}|^{\nu}|(q)) =$$

$$\sum_{\alpha_1=0}^{q_1+\nu l}\cdots\sum_{\alpha_m=0}^{q_m+\nu l}\left(\sum_{i_1=0}^{(\infty)}\cdots\sum_{i_m=0}^{(\infty)}((i)|\boldsymbol{A}|\,|(\alpha))\right)((\alpha)|\boldsymbol{A}|^{\nu}|(q)) \leq$$

$$M_{(q+\nu l)}\prod_{\beta=0}^{\nu-1} M_{(q+\beta l)} = \prod_{\beta=0}^{(\nu+1)-1} M_{(q+\beta l)}\,, \tag{33}$$

q. e. d.

Corollary 5.

For an m-dimensional column half matrix $\boldsymbol{A}$, *characterized by* (26), *we have*

$$(\Sigma\cdot|\boldsymbol{A}|^{\nu}|(q)) \leq (M_{(q)})^{\nu} \qquad \text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \quad ,, \quad ,, \quad \nu = 0, 1, 2, \ldots \tag{34}$$

For $\nu = 0$ (34) is, obviously, true and for $\nu > 0$ it follows immediately from (32) putting $l = 0$.

We can now generalize theorem 1 § 7.3 to the following theorem.

Theorem 3.

An arbitrary m-dimensional half matrix $\boldsymbol{A}(t)$ is absolutely exponentiable in each interval (x_0, x) and the exponential is again a half matrix. Furthermore, it is a continuous matrix function and has absolutely convergent row, respectively column, sums.

Firstly, it follows from corollary 3 that if the exponential exist, it is a half matrix. The rest of the theorem follows, using (30) and (34) respectively, by exactly the same argument as theorem 1 § 7.3, $\boldsymbol{K}$ being again given by (7.1.5). For a *row* half matrix we now only have, instead of (7.3.33),

$$|((p)|\exp[\boldsymbol{A}(t)(t-x_0)]\cdot\boldsymbol{\Sigma})| \leq \sum_{j_1=0}^{p_1}\cdots\sum_{j_m=0}^{p_m}\sum_{\nu=0}^{\infty}((p)|\boldsymbol{K}^\nu|(j))\frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}\left(\sum_{j_1=0}^{p_1}\cdots\sum_{j_m=0}^{p_m}((p)|\boldsymbol{K}^\nu|(j))\right)\frac{|x-x_0|^\nu}{\nu!} \leq \sum_{\nu=0}^{\infty}({}_{(p)}M)^\nu\frac{|x-x_0|^\nu}{\nu!} =$$

$$\exp\left[{}_{(p)}M|x-x_0|\right] < \infty \qquad \begin{array}{l}\text{for all } p_1, \ldots, p_m = 0, 1, 2, \ldots \\ \text{,, \;\; ,, \;\; intervals } (x_0, x) \\ \text{,, \;\; ,, \;\; } t \text{ in } x_0 \leq t \leq x.\end{array} \tag{35}$$

For a *column* half matrix we have, instead of (7.3.34),

$$|(\boldsymbol{\Sigma}\cdot\exp[\boldsymbol{A}(t)(t-x_0)]|(q))| \leq \sum_{i_1=0}^{q_1}\cdots\sum_{i_m=0}^{q_m}\sum_{\nu=0}^{\infty}((i)|\boldsymbol{K}^\nu|(q))\frac{|x-x_0|^\nu}{\nu!} =$$

$$\sum_{\nu=0}^{\infty}\left(\sum_{i_1=0}^{q_1}\cdots\sum_{i_m=0}^{q_m}((i)|\boldsymbol{K}^\nu|(q))\right)\frac{|x-x_0|^\nu}{\nu!} \leq \sum_{\nu=0}^{\infty}(M_{(q)})^\nu\frac{|x-x_0|^\nu}{\nu!} =$$

$$\exp\left[M_{(q)}|x-x_0|\right] < \infty \qquad \begin{array}{l}\text{for all } q_1, \ldots, q_m = 0, 1, 2, \ldots \\ \text{,, \;\; ,, \;\; intervals } (x_0, x) \\ \text{,, \;\; ,, \;\; } t \text{ in } x_0 \leq t \leq x.\end{array} \tag{36}$$

(35) and (36) prove the theorem.

Finally, we shall generalize theorem 1 § 7.4 to the following theorem which is the most important theorem for our applications of the theory (cf. § 4.8). In the condition (7.4.2) we only have to substitute q by a *linear* expression in $q_1, \ldots, q_m$, i. e.

$$(\boldsymbol{\Sigma}\cdot|\boldsymbol{A}(t)||(q)) \leq f(t)\cdot\sum_{i=1}^{m}c_iq_i \quad \text{for all } (q_1, \ldots, q_m) \neq (0, \ldots, 0) \tag{37}$$

in which $c_1, \ldots, c_m$ are arbitrary non-negative constants.

Theorem 4.

An arbitrary m-dimensional column semi-diagonal matrix $\boldsymbol{A}(t)$, characterized by (16), *which satisfies* (37) *is absolutely exponentiable in each interval* (x_0, x) *for which*

$$|x - x_0| < \frac{1}{\left(\sum_{i=1}^{m} c_i\right) lC} \tag{38}$$

with

$$C = \max_{x_0 \leq t \leq x} f(t)\ . \tag{39}$$

Furthermore, the exponential is continuous and has absolutely convergent column sums.

This theorem follows from (32), using (37) and (39), by exactly the same argument as theorem 1 § 7.4. Instead of (7.4.5) we now only have

$$\left|\left(\sum \cdot \exp\left[\boldsymbol{A}(t)(t-x_0)\right]|(q)\right)\right| \leq \sum_{i_1=0}^{\infty} \cdots \sum_{i_m=0}^{\infty} \sum_{\nu=0}^{\infty} \left((i)|\boldsymbol{K}^{\nu}|(q)\right) \frac{|x-x_0|^{\nu}}{\nu!} =$$

$$\sum_{\nu=0}^{\infty} \left(\sum_{i_1=0}^{(\infty)} \cdots \sum_{i_m=0}^{(\infty)} \left((i)|\boldsymbol{K}^{\nu}|(q)\right)\right) \frac{|x-x_0|^{\nu}}{\nu!} \leq \sum_{\nu=0}^{\infty} \left(\prod_{\beta=0}^{\nu-1} M_{(q+\beta l)} \right) \frac{|x-x_0|^{\nu}}{\nu!} \leq$$

$$\sum_{\nu=0}^{\infty} M_{(q)} \prod_{\beta=1}^{\nu-1} \left(\sum_{i=1}^{m} c_i(q_i+\beta l)\right) C^{\nu-1} \frac{|x-x_0|^{\nu}}{\nu!} = \sum_{\nu=0}^{\infty} u_{\nu} \tag{40}$$

for all $q_1, \ldots, q_m = 0, 1, 2, \ldots$
,, ,, $l = 0, 1, 2, \ldots$
,, ,, t in $x_0 \leq t \leq x$.

The series (40) is, however, convergent for all intervals satisfying (38) because

$$\frac{u_{\nu+1}}{u_{\nu}} = \frac{\sum_{i=1}^{m} c_i(q_i+\nu l)}{\nu+1} C|x-x_0| \underset{\nu\to\infty}{\to} \left(\sum_{i=1}^{m} c_i\right) lC|x-x_0| < 1, \tag{41}$$

q. e. d.

§ 7.6. Finally, we shall briefly summarize the definitions and theorems of chap. 7.

An arbitrary one- or m-dimensional matrix $\boldsymbol{A}(t)$ is said to be **absolutely exponentiable in an interval** (x_0, x) if the matrix function

$$\exp\left[\boldsymbol{K}(x-x_0)\right] = \sum_{\nu=0}^{\infty} \boldsymbol{K}^{\nu} \frac{(x-x_0)^{\nu}}{\nu!} \tag{7.1.7}$$

exist, when

$$\boldsymbol{K} = \max_{x_0 \leq t \leq x} |\boldsymbol{A}(t)| . \tag{7.1.5}$$

For the numerical row and column sums the following symbols have been introduced

$$((p)||\boldsymbol{A}|\cdot\Sigma) = \sum_{j_1=0}^{\infty}\dots\sum_{j_m=0}^{\infty}((p)||\boldsymbol{A}||(j)) \quad \text{for all } p_1,\dots,p_m = 0,1,2,\dots$$

(*row sum*) (2.1.10) and (3.1.10)

and

$$(\Sigma\cdot|\boldsymbol{A}||(q)) = \sum_{i_1=0}^{\infty}\dots\sum_{i_m=0}^{\infty}((i)||\boldsymbol{A}||(q)) \quad \text{for all } q_1,\dots,q_m = 0,1,2,\dots .$$

(*column sum*) (2.1.10) and (3.1.10)

A one- or m-dimensional matrix $\boldsymbol{A}$ is said to be **bounded** if a constant M exist so that

$$0 \leq (\Sigma\cdot|\boldsymbol{A}||(q)) \leq M < \infty \quad \text{for all } q_1,\dots,q_m = 0,1,2,\dots .$$

(7.1.12) and (7.5.5)

As a generalization of the Kronecker δ-symbol the following symbol is introduced

$$1(0 \leq \alpha \leq a) = \begin{cases} 1 & \text{for } 0 \leq \alpha \leq a \\ 0 & \text{,, } \alpha < 0 \text{ and } \alpha > a . \end{cases} \tag{7.2.3}$$

A one- or m-dimensional matrix $\boldsymbol{A}$ is said to be **row** (or **column**) **semi-finite** (generic term: **half semi-finite**) if for each fixed p- (or q-) point there is only a finite number of non-vanishing elements, i. e.

$$((p)|\boldsymbol{A}|(q)) = A((p),(q)) \prod_{i=1}^{m} 1(0 \leq q_i \leq {}_{(p)}n_i(\boldsymbol{A}))$$

(7.2.4) and (7.5.6)

when $\boldsymbol{A}$ is *row* semi-finite, and

$$((p)|\boldsymbol{A}|(q)) = A((p),(q)) \prod_{i=1}^{m} A(0 \leq p_i \leq n_i(\boldsymbol{A})_{(q)})$$

when $\boldsymbol{A}$ is *column* semi-finite. (7.2.2) and (7.5.7)

If both relations are satisfied, $\boldsymbol{A}$ is said to be semi-finite.

A one- or m-dimensional matrix $\boldsymbol{A}$ is said to be **row** (or **column**) **finite of order** N (generic term: **half finite of order** N) if it is half semi-finite and

$${}_{(p)}n_i(\boldsymbol{A}) \leq N \quad (\textit{row} \text{ finite}) \qquad (7.2.11) \text{ and } (7.5.13)$$

or

$$n_i(\boldsymbol{A})_{(q)} \leq N \quad (\textit{column} \text{ finite}) \qquad (7.2.12) \text{ and } (7.5.14)$$

respectively. If both relations are satisfied, $\boldsymbol{A}$ is said to be **finite of order** N.

A one- or m-dimensional matrix $\boldsymbol{A}$ is said to be **row** (or **column**) **semi-diagonal** (generic term: **half semi-diagonal**) if it is half semi-finite and an integer l exist so that

$${}_{(p)}n_i(\boldsymbol{A}) \leq p_i + l \quad (\textit{row} \text{ semi-diagonal}) \qquad (7.3.1) \text{ and } (7.5.15)$$

or

$$n_i(\boldsymbol{A})_{(q)} \leq q_i + l \quad (\textit{column} \text{ semi-diagonal}) \qquad (7.3.2) \text{ and } (7.5.16)$$

respectively. If both relations are satisfied, $\boldsymbol{A}$ is said to be **semi-diagonal.**

A one- or m-dimensional matrix $\boldsymbol{A}$ is said to be a **row** (or **column**) **half matrix** if $\boldsymbol{A}$ is half semi-diagonal with $l = 0$. (If $\boldsymbol{A}$ is both a row *and* a column half matrix, it is simply a diagonal matrix.)

Theorem I (1 § 7.1 and 1 § 7.5).

An arbitrary one- or m-dimensional bounded matrix $\boldsymbol{A}(t)$ *is absolutely exponentiable in each interval if* $\boldsymbol{K}$ *given in* (7.1.5) *is also bounded. The exponential, being again bounded, is, furthermore, continuous.*

In this case we shall say that matrix $\boldsymbol{A}$ is of **type I**. Thus Feller's probability operators (cf. (2.1.14)) are of type I.

Theorem II (1 § 7.2 and 2 § 7.5).

An arbitrary one- or m-dimensional half finite matrix $\boldsymbol{A}(t)$ *of order* N *is absolutely exponentiable in each interval and the exponential is again half finite of order* N. *Furthermore, the exponential is continuous and has absolutely convergent row, respectively column, sums.*

In this case we shall say that matrix $\boldsymbol{A}$ is of **type II**. Thus Kolmogoroff's probability operators (cf. § 2.8) are of type II.

Theorem III (1 § 7.3 and 3 § 7.5).

An arbitrary one- or m-dimensional half matrix $\boldsymbol{A}(t)$ *is absolutely exponentiable in each interval and the exponential is again half. Furthermore, the exponential is continuous and has absolutely convergent row, respectively column, sums.*

In this case we shall say that matrix $\boldsymbol{A}$ is of **type III**. Thus Lundberg's probability operators (cf. (2.1.15)) are of type III.

Theorem IV (1 § 7.4 and 4 § 7.5).

An arbitrary one- or m-dimensional column semi-diagonal matrix $\boldsymbol{A}(t)$ *whose numerical column sums are bounded by the relation*

$$\left(\Sigma \cdot |\boldsymbol{A}(t)|(q)\right) \leq f(t) \cdot \sum_{i=1}^{m} c_i q_i \quad \text{for all } (q_1, \ldots, q_m) \neq (0, \ldots, 0) \qquad (7.4.2) \text{ and } (7.5.37)$$

is absolutely exponentiable in each interval (x_0, x) *for which*

$$|x - x_0| < \frac{1}{\left(\sum_{i=1}^{m} c_i\right) e C} \qquad (7.4.3) \text{ and } (7.5.38)$$

with

$$C = \max_{x_0 \leq t \leq x} f(t) . \qquad (7.4.4) \text{ and } (7.5.39)$$

Furthermore, the exponential is continuous and has absolutely convergent column sums.

In this case we shall say that matrix $\boldsymbol{A}$ is of **type IV**. Thus the probability operators in chap. 4 are of type IV.

CHAPTER 8.

On the numerical Computation of $\varphi(x) = \int_0^x \exp[t^2]dt.$

§ 8.1. In chaps. 4 and 6 we have needed the numerical values of the function

$$\varphi(x) = -\varphi(-x) = \int_0^x \exp[t^2]dt\,. \tag{1}$$

This function is tabulated to four places in JAHNKE-EMDE[1]) in the interval $x = 0{\cdot}00-(0{\cdot}01)-1{\cdot}99$ on the basis of the corresponding 6-place values given by DAWSON[2]). We cannot, however, confine ourselves to this interval and shall, therefore, in this paragraph, extending the calculations of DAWSON, give $\varphi(x)$ to 4 places in the interval $x = 2{\cdot}00-(0{\cdot}01)-10{\cdot}00$.

Now the main contribution to the integral in (1) comes from the values of $t \sim x$. We therefore write the integral in such a way that we can expand in a rapidly convergent series from the point $t = x$

$$\varphi(x) = \int_0^x \exp[t^2]dt = \int_0^x \exp\left[(x-u)^2\right]du =$$

$$\exp[x^2]\int_0^x \exp[u^2 - 2xu]du = \frac{\exp[x^2]}{2x}\int_0^{2x^2} \exp\left[\frac{t^2}{4x^2} - t\right]dt =$$

$$\frac{\exp[x^2]}{2x} f(x) \tag{2}$$

in which we have put

$$f(x) = \frac{\varphi(x)}{\dfrac{\exp[x^2]}{2x}} = \int_0^{2x^2} \exp\left[\frac{t^2}{4x^2} - t\right]dt \underset{x\to\infty}{\to} \int_0^\infty e^{-t}dt = 1\,. \tag{3}$$

[1]) JAHNKE and EMDE (1933). Cf. p. 106.

[2]) DAWSON (1898).

Table 23. $f(x) = \dfrac{\psi(x)}{\dfrac{\exp[x^2]}{2x}} = 1 + \dfrac{1}{2x^2} + \dfrac{1 \cdot 3}{(2x^2)^2} + \dfrac{1 \cdot 3 \cdot 5}{(2x^2)^3} + \ldots$ *for* $x \gg 1$.

x	0	I	2	3	4	5	6	7	8	9
2·0	1·205	03	01	*98	96	94	92	90	87	85
2·1	1·183	81	79	77	75	73	71	69	68	66
2·2	1·164	62	60	58	56	54	52	50	49	47
2·3	1·145	43	42	40	39	37	36	34	33	31
2·4	1·130	29	27	26	24	23	22	20	19	17
2·5	1·116	15	14	12	11	10	09	08	06	05
2·6	1·104	03	02	00	*99	98	97	96	95	94
2·7	1·093	92	91	90	89	88	87	86	86	85
2·8	1·084	83	82	82	81	80	79	78	78	77
2·9	1·076	75	74	74	74	72	71	71	70	70
3·0	1·069	68	68	67	67	66	66	65	65	64
3·1	1·064	64	63	63	62	62	62	61	61	60
3·2	1·060	59	59	58	58	57	57	56	56	55
3·3	1·055	55	54	54	53	53	53	52	52	51
3·4	1·051	51	50	50	49	49	49	48	48	47
3·5	1·047	47	47	46	46	46	46	45	45	44
3·6	1·044	44	44	44	43	43	43	42	42	41
3·7	1·041	41	41	40	40	40	40	39	39	38
3·8	1·038	38	38	38	37	37	37	37	36	36
3·9	1·036	36	36	35	35	35	35	35	35	34
4·	1·034	32	31	29	28	26	25	24	23	22
5·	1·0210	203	196	188	180	173	167	162	156	150
6·	1·0145	141	137	132	128	124	120	117	113	109
7·	1·0105	102	100	098	095	092	090	087	084	082
8·	1·0080	078	076	074	072	070	069	068	066	064
9·	1·0063	061	060	060	059	058	056	055	054	052
10·	1·0051									

In $f(x)$ we now expand $\exp\left|\dfrac{t^2}{4x^2}\right|$ in Taylor's series and thus find

$$\psi(x) = \frac{\exp[x^2]}{2x} \sum_{j=0}^{\infty} \frac{1}{j!} \left(\frac{1}{2x}\right)^{2j} \int_0^{2x^2} t^{2j} e^{-t} dt =$$

$$\frac{\exp[x^2]}{2x} \left(\frac{1}{0!} (1 - \exp[-2x^2]) + \frac{1}{1!} \left(\frac{1}{2x}\right)^2 (2 - \exp[-2x^2](4x^4 + 4x^2 + 2)) + \right.$$

$$\left. \frac{1}{2!} \left(\frac{1}{2x}\right)^4 (24 - \exp[-2x^2](16x^8 + 32x^6 + 48x^4 + 48x^2 + 24)) + \ldots \right). \quad (4)$$

Table 24. $\psi(x) = \int_0^x \exp[t^2]dt$ [1]

x	0	1	2	3	4	5	6	7	8	9
2·0	$10 \times 1·645$	1·701	1·759	1·818	1·881	1·947	2·015	2·087	2·159	2·237
2·1	$10 \times 2·317$	2·401	2·489	2·581	2·676	2·776	2·880	2·988	3·104	3·222
2·2	$10 \times 3·345$	3·474	3·610	3·750	3·898	4·051	4·212	4·380	4·560	4·744
2·3	$10 \times 4·937$	5·139	5·354	5·575	5·812	6·054	6·314	6·580	6·865	7·158
2·4	$10 \times 7·470$	7·800	8·137	8·499	8·871	9·269	9·687	10·12	10·58	11·05
2·5	$10^2 \times 1·156$	1·210	1·266	1·326	1·386	1·451	1·520	1·592	1·667	1·747
2·6	$10^2 \times 1·830$	1·920	2·014	2·110	2·214	2·324	2·439	2·560	2·689	2·824
2·7	$10^2 \times 2·966$	3·117	3·276	3·443	3·620	3·807	4·005	4·213	4·438	4·671
2·8	$10^2 \times 4·916$	5·178	5·453	5·750	6·059	6·384	6·729	7·095	7·489	7·899
2·9	$10^2 \times 8·333$	8·793	9·280	9·806	10·36	10·93	11·55	12·21	12·91	13·66
3·0	$10^3 \times 1·444$	1·527	1·616	1·710	1·811	1·916	2·030	2·150	2·279	2·413
3·1	$10^3 \times 2·559$	2·715	2·877	3·053	3·237	3·436	3·648	3·871	4·112	4·364
3·2	$10^3 \times 4·639$	4·925	5·235	5·562	5·915	6·286	6·688	7·111	7·570	8·051
3·3	$10^3 \times 8·574$	9·132	9·719	10·36	11·03	11·76	12·53	13·35	14·24	15·18
3·4	$10^4 \times 1·620$	1·729	1·844	1·969	2·101	2·244	2·398	2·560	2·737	2·926
3·5	$10^4 \times 3·127$	3·343	3·576	3·827	4·092	4·380	4·690	5·017	5·374	5·752
3·6	$10^4 \times 6·163$	6·612	7·082	7·593	8·135	8·727	9·363	10·04	10·77	11·56
3·7	$10^5 \times 1·244$	1·333	1·432	1·537	1·651	1·775	1·909	2·050	2·205	2·370
3·8	$10^5 \times 2·550$	2·745	2·955	3·181	3·423	3·686	3·972	4·280	4·608	4·968
3·9	$10^5 \times 5·356$	5·777	6·231	6·717	7·248	7·823	8·446	9·120	9·850	10·63
4·0	$10^6 \times 1·149$	1·241	1·342	1·451	1·567	1·695	1·833	1·984	2·147	2·322
4·1	$10^6 \times 2·513$	2·722	2·950	3·195	3·462	3·748	4·064	4·406	4·778	5·182
4·2	$10^6 \times 5·623$	6·101	6·617	7·182	7·798	8·470	9·200	9·996	10·86	11·81
4·3	$10^7 \times 1·282$	1·395	1·517	1·650	1·795	1·954	2·124	2·313	2·519	2·743
4·4	$10^7 \times 2·988$	3·256	3·548	3·865	4·214	4·595	5·012	5·468	5·967	6·512
4·5	$10^7 \times 7·102$	7·754	8·468	9·250	10·11	11·04	12·07	13·18	14·41	15·76
4·6	$10^8 \times 1·724$	1·887	2·064	2·260	2·474	2·699	2·964	3·247	3·557	3·898
4·7	$10^8 \times 4·273$	4·685	5·137	5·633	6·175	6·775	7·435	8·162	8·961	9·800
4·8	$10^9 \times 1·081$	1·187	1·303	1·433	1·575	1·731	1·904	2·094	2·304	2·535
4·9	$10^9 \times 2·790$	3·072	3·379	3·721	4·098	4·515	4·976	5·484	6·046	6·666
5·0	$10^9 \times 7·352$	8·109	8·947	9·874	10·90	12·03	13·27	14·66	16·19	17·89
5·1	$10^{10} \times 1·978$	2·185	2·416	2·671	2·955	3·269	3·616	4·002	4·430	4·905
5·2	$10^{10} \times 5·430$	6·015	6·664	7·377	8·175	9·062	10·05	11·14	12·36	13·71
5·3	$10^{11} \times 1·521$	1·688	1·874	2·079	2·309	2·564	2·840	3·166	3·518	3·911
5·4	$10^{11} \times 4·349$	4·836	5·379	5·985	6·660	7·412	8·251	9·179	10·22	11·39
5·5	$10^{12} \times 1·260$	1·414	1·576	1·757	1·956	2·184	2·428	2·719	3·034	3·387
5·6	$10^{12} \times 3·780$	4·222	4·715	5·267	5·885	6·570	7·344	8·210	9·181	10·27
5·7	$10^{13} \times 1·149$	1·285	1·439	1·610	1·803	2·019	2·261	2·533	2·838	3·181
5·8	$10^{13} \times 3·564$	3·997	4·483	5·023	5·635	6·323	7·095	7·966	8·944	10·04
5·9	$10^{14} \times 1·131$	1·268	1·424	1·601	1·799	2·023	2·275	2·559	2·879	3·240

[1]) For $0·00 \leq x \leq 2·00$ the values of $\psi(x)$ are given i JAHNKE-EMDE (2. ed.) p. 106.

Table 24 cont.

x	0	1	2	3	4	5	6	7	8	9
6·0	$10^{14}\times 3{\cdot}644$	4·101	4·616	5·200	5·855	6·600	7·437	8·383	9·450	10·66
6·1	$10^{15}\times 1{\cdot}202$	1·356	1·529	1·726	1·948	2·199	2·483	2·804	3·168	3·579
6·2	$10^{15}\times 4{\cdot}044$	4·571	5·169	5·845	6·610	7·478	8·454	9·565	10·83	12·26
6·3	$10^{16}\times 1{\cdot}388$	1·572	1·781	2·018	2·287	2·592	2·939	3·333	3·780	4·288
6·4	$10^{16}\times 4{\cdot}864$	5·522	6·268	7·116	8·080	9·178	10·43	11·85	13·45	15·29
6·5	$10^{17}\times 1{\cdot}739$	1·977	2·249	2·558	2·911	3·313	3·771	4·294	4·890	5·571
6·6	$10^{17}\times 6{\cdot}346$	7·214	8·241	9·394	10·71	12·22	13·93	15·90	18·14	20·70
6·7	$10^{18}\times 2{\cdot}363$	2·699	3·082	3·520	4·022	4·596	5·253	5·999	6·860	7·845
6·8	$10^{18}\times 8{\cdot}977$	10·27	11·75	13·45	15·40	17·63	20·19	23·13	26·38	30·36
6·9	$10^{19}\times 3{\cdot}480$	3·990	4·575	5·247	6·019	6·907	7·926	9·098	10·44	11·99
7·0	$10^{20}\times 1{\cdot}377$	1·581	1·816	2·097	2·399	2·758	3·172	3·648	4·196	4·826
7·1	$10^{20}\times 5{\cdot}558$	6·396	7·364	8·480	9·767	11·25	12·96	14·94	17·27	19·86
7·2	$10^{21}\times 2{\cdot}290$	2·641	3·047	3·515	4·057	4·683	5·407	6·244	7·212	8·332
7·3	$10^{21}\times 9{\cdot}626$	11·13	12·86	14·87	17·20	19·89	23·01	26·63	30·82	35·68
7·4	$10^{22}\times 4{\cdot}129$	4·779	5·535	6·413	7·431	8·612	9·984	11·58	13·43	15·57
7·5	$10^{23}\times 1{\cdot}807$	2·096	2·433	2·823	3·277	3·809	4·424	5·140	5·973	6·942
7·6	$10^{23}\times 8{\cdot}070$	9·384	10·91	12·69	14·77	17·19	20·00	23·29	27·12	31·51
7·7	$10^{24}\times 3{\cdot}678$	4·286	4·995	5·822	6·787	7·914	9·229	10·77	12·56	14·64
7·8	$10^{25}\times 1{\cdot}710$	1·995	2·330	2·721	3·179	3·714	4·341	5·080	5·931	6·936
7·9	$10^{25}\times 8{\cdot}114$	9·491	11:10	12·99	15·21	17·81	20·85	24·42	28·61	33·52
8·0	$10^{26}\times 3{\cdot}928$	4·604	5·398	6·330	7·425	8·710	10·22	11·99	14·08	16·53
8·1	$10^{27}\times 1{\cdot}941$	2·280	2·678	3·147	3·698	4·347	5·111	6·010	7·069	8·316
8·2	$10^{27}\times 9{\cdot}786$	11·52	13·56	15·95	18·78	22·12	26·06	30·70	36·18	42·65
8·3	$10^{28}\times 5{\cdot}031$	5·930	6·995	8·252	9·737	11·49	13·57	16·02	18·92	22·34
8·4	$10^{29}\times 2{\cdot}640$	3·119	3·686	4·358	5·152	6·093	7·207	8·527	10·09	11·94
8·5	$10^{30}\times 1{\cdot}413$	1·674	1·982	2·348	2·782	3·296	3·907	4·632	5·492	6·513
8·6	$10^{30}\times 7{\cdot}725$	9·165	10·88	12·91	15·33	18·20	21·61	25·67	30·50	36·24
8·7	$10^{31}\times 4{\cdot}307$	5·121	6·089	7·242	8·614	10·25	12·20	14·47	17·28	20·58
8·8	$10^{32}\times 2{\cdot}450$	2·919	3·478	4·145	4·940	5·890	7·016	8·367	9·981	11·91
8·9	$10^{33}\times 1{\cdot}422$	1·696	2·025	2·418	2·888	3·450	4·122	4·926	5·888	7·039
9·0	$10^{33}\times 8{\cdot}419$	10·07	12·04	14·41	17·24	20·64	24·71	29·59	35·43	42·46
9·1	$10^{34}\times 5{\cdot}087$	6·096	7·307	8·761	10·51	12·60	15·12	18·14	21·77	26·13
9·2	$10^{35}\times 3{\cdot}137$	3·767	4·524	5·435	6·530	7·848	9·433	11·34	13·64	16·40
9·3	$10^{36}\times 1{\cdot}973$	2·375	2·858	3·440	4·142	4·987	6·007	7·237	8·720	10·51
9·4	$10^{37}\times 1{\cdot}267$	1·527	1·842	2·222	2·680	3·234	3·903	4·711	5·688	6·869
9·5	$10^{37}\times 8{\cdot}295$	10·02	12·11	14·64	17·70	21·40	25·87	31·29	37·86	45·81
9·6	$10^{38}\times 5{\cdot}543$	6·712	8·127	9·842	11·92	14·44	17·50	21·21	25·71	31·17
9·7	$10^{39}\times 3{\cdot}779$	4·586	5·564	6·744	8·186	9·937	12·07	14·65	17·80	21·62
9·8	$10^{40}\times 2{\cdot}629$	3·194	3·884	4·720	5·740	6·983	8·496	10·34	12·58	15·32
9·9	$10^{41}\times 1{\cdot}866$	2·261	2·754	3·356	4·089	5·007	6·104	7·442	9·076	11·07
10·0	$10^{42}\times 1{\cdot}351$									

For $x > 10$ we have correctly to 5 places $\psi(x) = \dfrac{\exp[x^2]}{2x}\left(1+\dfrac{1}{2x^2}+\dfrac{3}{4x^4}\right)$.

From (4) we see that if $x \gg 1$ we may, within our limits of accuracy, replace the upper limit in the integrals, $2x^2$, by ∞. We thus find from (4) the asymptotic formula

$$\psi(x) = \frac{\exp[x^2]}{2x}\left(1 + \frac{1}{2x^2} + \frac{1\cdot 3}{(2x^2)^2} + \frac{1\cdot 3\cdot 5}{(2x^2)^3} + \ldots\right) \quad x \gg 1\,. \qquad (5)$$

From Dawson's table we find e. g. for $x = 2$ $\psi(2) = 16{\cdot}45$, i. e. $f(2) = \frac{16{\cdot}45}{13{\cdot}65} = 1{\cdot}205$ to be compared with the values given in (5)

$$f(2) = 1 + \frac{1}{2\cdot 4} = 1{\cdot}125,\quad 1 + \frac{1}{2\cdot 4} + \frac{3}{4\cdot 16} = 1{\cdot}172,$$

$$1 + \frac{1}{2\cdot 4} + \frac{3}{4\cdot 16} + \frac{15}{8\cdot 64} = 1{\cdot}201\,.$$

For values below $x = 5$ DAWSON has given the formula

$$\psi(x) = \frac{1}{2\sqrt{\pi}}\left(\psi_1(x) + e^{-\frac{1}{4}}\psi_2(x)\right) + \varepsilon(x) \qquad (6)$$

in which

$$\left.\begin{aligned}\psi_1(x) &= x + \sum_{n=1}^{\infty}\frac{1}{n}\,\frac{\sinh(2nx)}{\exp[n^2]}\\ \psi_2(x) &= \sum_{n=0}^{\infty}\frac{2}{2n+1}\,\frac{\sinh\left((2n+1)x\right)}{\exp\left[n(n+1)\right]}\,.\end{aligned}\right\} \qquad (7)$$

The error in (6) does not exceed

$$|\varepsilon(x)| < \int_0^x \exp[t^2 - 4\pi^2]\,|\cos 4\pi t|\,dt < \frac{\exp[-4\pi^2 + x^2]}{4\pi} < 5\times 10^{-8}$$

$$\text{for } x \leqq 5\,. \qquad (8)$$

By means of Dawson's table and formulae we have now computed $f(x)$ for $x = 1{\cdot}0$, $1{\cdot}2$, $1{\cdot}4$, $1{\cdot}5$, $1{\cdot}6$, $1{\cdot}7$, $1{\cdot}8$, $1{\cdot}85$, $1{\cdot}9$, $1{\cdot}95$, $2{\cdot}0$, $2{\cdot}2$, $2{\cdot}4$, $2{\cdot}6$, $2{\cdot}8$, $3{\cdot}0$, $3{\cdot}5$, $4{\cdot}0$ and $5{\cdot}0$. From (5) we have next computed $f(x)$ for $x = 6$, 7, 8, 9 and 10. The figures so obtained were plotted on millimetre-paper, as shown in fig. 54 and the values of $f(x)$ were read off for every $0{\cdot}1$ in the interval $x = 2{\cdot}0$–$10{\cdot}0$ and tested by means of the differences. Next we computed by subdividing the differences the values of $f(x)$ in the interval $x = 2{\cdot}00$–$4{\cdot}00$ for every $0{\cdot}01$. The results are given in table 23. We estimate that
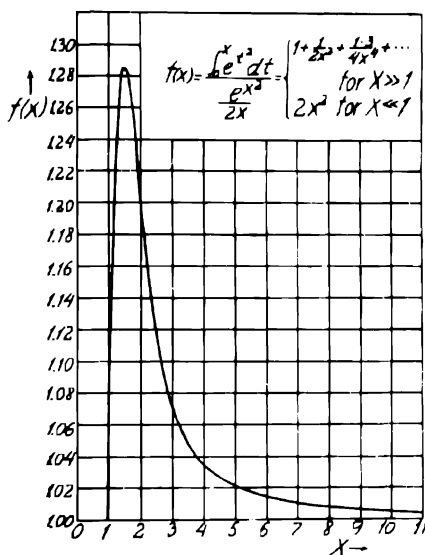
FIG. 54.

the figures are correct within 1 or 2 units in the last figure. In table 24 we finally give the corresponding values of $\varphi(x)$ computed from (2).

SUMMARY.

The purpose of the present paper is to investigate in further detail the *fluctuation problem* in the theory of cosmic radiation. This radiation is, as reviewed in chap. 5, divided into a *soft component*, consisting of positive and negative electrons together with electromagnetic radiation quanta, photons, and a *hard component*, consisting essentially of mesons. By the passage through matter the soft component gives rise to the formation of *multiplication* or *cascade showers* consisting of secondary electrons and photons. These particles are formed by the successive transformation of electrons into photons and vice versa by a process which in a certain way is analogous to a *biological process* in which one generation of individuals produces a new one by propagation or is effaced by death. It is the theory of this shower phenomenon, being fundamentally of a statistical nature, which is treated in the present paper on the basis of the fundamental papers of Bhabha and Heitler (1937) and Carlson and Oppenheimer (1937) together with the previous papers of the author (1938 and 1940). As outlined in chap. 5 this theory allows of a calculation of the mean number, $\overline{N}$, of secondary electrons which are produced by a primary electron or photon with given energy in a layer of matter of given thickness. These showers may be investigated experimentally by means of Geiger-Müller counters in coincidence coupling. The resulting curves, the *Rossi curves*, give essentially the probabilities, as functions of the thickness of the layer penetrated, of obtaining a shower containing at least one, two, three electrons and so on, depending on the experimental arrangement. As a consequence we cannot confine ourselves to the consideration of the mean numbers, $\overline{N}$, but must evaluate direct the probabilities themselves of obtaining showers with only one, two, three electrons and so on. This *fluctuation problem* cannot, however, be solved direct due to the complicated nature of the problem. It is necessary to be

content with the construction of *idealized models* of the real multiplication process. In chap. 4 we discuss the models previously suggested together with the generalizations of the author. In chap. 6 we describe how it is possible to apply, on the basis of the results mentioned in chaps. 4 and 5, the models to the theoretical calculation of the experimental Rossi curves. Fortunately enough, it turns out that the calculation does not depend significantly on the fluctuation problem as the resulting curves vary little from model to model. It is pointed out that the reason for this fact lies in the necessity of averaging the probabilities in question over the primary *energy spectra*. This operation causes the fluctuations around the mean electron numbers to be reduced considerably as the fluctuations are, of course, the larger the larger the primary energy. Finally, the calculated Rossi curves are compared with the rather scanty experiments available. The agreement is particularly good, indeed surprisingly good taking into account the long series of approximations which have been necessary. It is stressed that an *experimental* separation of the various constituents of cosmic radiation is essential for the comparison between theory and experiments and that *further experiments are very desirable.*

The models just mentioned represent special cases of the *stochastic processes* which in later years have been the object of much interest in the theory of probability. The stochastic processes have, namely, turned out to be extremely valuable for the discussion of statistical problems in a great variety of fields as, e. g., actuarial science, the theory of telephones, theoretical biology and, as discussed in part II of the present paper, also physics. The systematic theory of the stochastic processes is in the first place due to KOLMOGOROFF (1931) and FELLER (1937). LUNDBERG (1940) has generalized the theory of the *discontinuous* stochastic processes to embrace also the processes met with in actuarial problems. It turned out, however, that this generalization was not sufficiently wide to cover also the processes needed in the theory of cosmic radiation. Notwithstanding the fact that physical intuition, of course, ensures a legitimate performance of the calculations required for the *special* processes considered and that the mathematical theory, therefore, is rather a luxury, it is, nevertheless, of a certain interest, at any rate from a purely mathematical point of view, to give a systematic generalization of the theory in order to embrace also the processes in question. It is the purpose of the first part of the paper to give such

a generalization. As this part gives only the theoretical background for the second part of the paper, and as presumably only few readers will be interested in both the mathematical theory and the physical applications, either part has been written to form a complete entity. In chap. 1 we give a survey of the general theory of stochastic processes. In chap. 2 we give the generalization of the author of the theory of one-dimensional processes, i. e. processes in which only one stochastic variable enters. The mathematical instrument necessary for this purpose consists in the solution of *infinite* systems of simultaneous, linear differential equations of the first order. It is shown that, introducing the convenient *matrix symbolism*, the usual theory, both the method of solving by means of *iteration* as well as the representation of the solution in the form of a *product-integral*, may be immediately generalized from finite to infinite systems of equations when only it is assumed that the operator-matrix of the system has a purely algebraic property for which the term *exponentiability* is proposed. In an appendix, chap. 7, this notion is discussed in greater detail as well as the conditions for a, one- or multi-dimensional, matrix to be exponentiable. It is shown that presumably all processes met with in the practical applications of the theory are covered by the generalization in question. (Cf. § 7.6 in which all the definitions and theorems regarding exponentiable matrices are summarized.) In chap. 2 we next show how the unique solution will automatically satisfy all the requirements, except one, which must be demanded of the solution in order that it shall represent a stochastic process. The only requirement, the fulfilment of which necessitates further conditions, is the claim that the entire 'probability mass' shall remain in the finite for all times, i. e. that the sum of the probabilities over all possible states shall remain a constant equal to one. The conditions of FELLER and LUNDBERG for the fulfilment of this claim are generalized. Moreover, the question of the existence of the moments is investigated. Finally, it is shown in chap. 3 that the results obtained in chap. 2 can be immediately generalized to multi-dimensional processes, i. e. processes in which more stochastic variables enter as is, e. g., the case in the processes used in the theory of cosmic radiation (chap. 4). This is simply due to the well-known fact that 'dimension' is an artificial conception from the point of view of the theory of manifolds.

DANSK RESUMÉ.

Formaalet med det foreliggende arbejde er en nærmere undersøgelse af *fluktuationsproblemet* indenfor teorien for den kosmiske straaling. Denne straaling deles, som omtalt i kap. 5, i en *blød komponent*, bestaaende af positive og negative elektroner samt elektromagnetiske straalingskvanter, fotoner, og en *haard komponent*, bestaaende væsentligst af mesoner. Den bløde komponent giver ved gennemgang gennem materie anledning til dannelsen af *multiplikations-* eller *kaskade-byger*, bestaaende af sekundære elektroner og fotoner. Disse dannes ved successive omdannelser af elektroner til fotoner og omvendt i en proces, som i en vis forstand er analog til en *biologisk proces*, hvor een generation af individer frembringer en ny ved formering eller udslettes ved død. Det er teorien for dette bygefænomen, der er af principiel statistisk natur, som behandles i det foreliggende arbejde paa grundlag af de fundamentale arbejder af BHABHA og HEITLER (1937) og CARLSON og OPPENHEIMER (1937), samt forfatterens tidligere arbejder (1938 og 1940). Som skitseret i kap. 5 tillader denne teori en beregning af det gennemsnitlige antal, $\bar{N}$, af sekundære elektroner, som frembringes af en primær elektron eller foton med given energi i et stoflag med given tykkelse. Med GEIGER-MÜLLER tællere i koincidensopstilling kan disse byger undersøges experimentelt. De resulterende kurver, *Rossikurverne*, angiver i det væsentligste sandsynlighederne, som funktioner af lagtykkelsen, for at faa en byge bestaaende af mindst en, to, tre elektroner o. s. v. alt efter den experimentelle opstilling. Man kan derfor ikke nøjes med at beregne det gennemsnitlige antal, $\bar{N}$, men maa direkte udregne selve sandsynlighederne for at faa byger med netop en, to, tre elektroner o. s. v. Dette *fluktuationsproblem* kan imidlertid ikke løses direkte paa grund af problemets komplicerede natur. Man maa nøjes med at opstille *idealiserede modeller* af den virkelige multiplikationsproces. I kap. 4 diskuteres de modeller, der tidligere har været foreslaaet, samt forfatterens

videreførelse heraf. I kap. 6 beskrives det, hvorledes man paa grundlag af de i kap. 4 og 5 angivne resultater kan anvende modellerne til en teoretisk beregning af de experimentelle Rossikurver. Det viser sig herved, at fluktuationsproblemet heldigvis ikke indgaar paa kritisk maade i beregningerne, idet de resulterende kurver kun varierer lidt fra model til model. Det paavises, at grunden hertil er, at man maa danne middelværdierne af de paagældende sandsynligheder over *energispektrene* af de primære partikler, hvilket bevirker, at fluktuationerne omkring de gennemsnitlige elektronantal nedsættes meget væsentligt, idet fluktuationerne naturligvis er desto større, jo større den primære energi er. Endelig sammenlignes de fundne Rossikurver med de temmelig faa foreliggende experimenter. Overensstemmelsen er særdeles god, ja nærmest forbløffende god, naar man betænker den lange række af approximationer, som det har været nødvendigt at foretage. Det fremhæves, at en *experimentel* adskillelse mellem den kosmiske straalings forskellige bestanddele er væsentlig for sammenligningen mellem teori og experimenter, og at *yderligere experimenter er særdeles ønskelige*.

De netop nævnte modeller repræsenterer specialtilfælde af de *stochastiske processer*, som i de senere aar har været genstand for stor interesse indenfor sandsynlighedsregningen. Stochastiske processer har nemlig vist sig at frembyde et særdeles værdifuldt hjælpemiddel ved diskussionen af statistiske problemer indenfor de mest forskellige omraader som for exempel aktuarvidenskab, telefonteori, teoretisk biologi og, som diskuteret i anden del af det foreliggende arbejde, ogsaa fysik. Den systematiske teori for de stochastiske processer skyldes først og fremmest KOLMOGOROFF (1931) og FELLER (1937). LUNDBERG (1940) har generaliseret teorien for de *diskontinuerte* stochastiske processer til ogsaa at omfatte de processer, man har brug for indenfor aktuarvidenskaben. Det viste sig imidlertid, at denne generalisation ikke var tilstrækkelig omfattende til ogsaa at dække de processer, man har brug for i teorien for den kosmiske straaling. Selvom den fysiske intuition selvfølgelig garanterer, at de foretagne regneoperationer er lovlige for de specielle processer, der betragtes, og den matematiske teori saaledes nærmest maa betragtes som en luxus, er det dog af en vis interesse, i hvert fald fra et rent matematisk synspunkt, at foretage en systematisk generalisation af teorien, saaledes at den ogsaa kommer til at omfatte de paagældende processer. Dette er formaalet med arbejdets første del. Da denne saaledes kun giver den teoretiske baggrund for

arbejdets anden del, og da formodentlig kun faa læsere vil være interesseret i baade den matematiske teori og de fysiske anvendelser, er begge dele skrevet saaledes, at de hver for sig udgør et afsluttet hele. I kap. 1 gives en oversigt over den almindelige teori for stochastiske processer. I kap. 2 følger forfatterens generalisation af teorien for en-dimensionale processer, d. v. s. processer, hvori der kun indgaar een stochastisk variabel. Det hertil nødvendige matematiske hjælpemiddel bestaar i løsningen af *uendelige* systemer af simultane, lineære differentialligninger af første orden. Det vises, at naar man indfører den bekvemme *matrixsymbolisme*, kan den sædvanlige teori, saavel metoden at løse ved *iteration* som løsningens fremstilling i form af et *produktintegral*, umiddelbart generaliseres fra endelige ligningssystemer til uendelige, naar det blot antages, at ligningssystemets operatormatrix har en rent algebraisk egenskab, for hvilken der foreslaas betegnelsen *exponentiabilitet*. I et tillæg, kap. 7, gives en nærmere diskussion af dette begreb samt af betingelserne for, at en en-dimensional eller fler-dimensional matrix er exponentiabel. Det vises, at formodentlig alle processer, man kan møde ved de praktiske anvendelser af teorien, omfattes af den foretagne generalisation. (Sml. § 7.6, i hvilken alle definitioner og sætninger vedrørende exponentiable matricer resumeres.) I kap. 2 vises det dernæst, hvorledes den entydige løsning netop automatisk opfylder alle de fordringer, paa nær een, som maa stilles til en løsning, for at den skal fremstille en stochastisk proces. Den eneste fordring, hvis opfyldelse nødvendiggør yderligere betingelser, er fordringen om, at hele »sandsynlighedsmassen« skal forblive i det endelige til alle tider, d. v. s. at summen af sandsynlighederne udstrakt over alle mulige tilstande skal forblive lig med 1. De af FELLER og LUNDBERG opstillede betingelser herfor generaliseres. Desuden undersøges spørgsmaalet om momenternes existens. Endelig vises det i kap. 3, at de i kap. 2 opnaaede resultater umiddelbart kan overføres til fler-dimensionale processer, d. v. s. processer, hvori der indgaar flere stochastiske variable, saaledes som det for exempel er tilfældet med de i teorien for den kosmiske straaling anvendte processer (kap. 4). Grunden hertil er simpelthen det velkendte forhold, at »dimension« er et kunstigt begreb set fra mængdelærens synspunkt.

LIST OF REFERENCES.

ARLEY 1938, Proc. Roy. Soc. A **168**, 519. (Quoted as A.)
— 1940, Fys. Tidsskr. **38**, 74. (In Danish.)
— and BUCH 1943, Sandsynlighedsregning med Anvendelse paa Statistik, Fejlteori og Udjævningsteori, 2. Udg. København. (Quoted as A & B.)
— and ERIKSEN 1940, Kgl. Dansk. Vid. Selsk. Skr. **XVII** no. **11**. (Quoted as A & E.)
— and HEITLER 1938, Nature **142**, 158.
AUGER, LEPRINCE-RINGUET et EHRENFEST jr. 1936, J. Phys. Radium **7**, 58.

BACHELIER 1900, Ann. l'École Norm. **17**, 21.
— 1912, Calcul des probabilités, Paris.
BHABHA 1938, Proc. Roy. Soc. A **164**, 257.
— and HEITLER 1937, Proc. Roy. Soc. A **159**, 432. (Quoted as B & H.)
BRADDICK 1939, Cosmic Rays and Mesotrons, Cambridge.
BRUINS 1941, Physica **8**, 988.

CARLSON and OPPENHEIMER 1937, Phys. Rev. **51**, 220.
CRAMÉR 1937, Random Variables and Probability Distributions, Cambridge.

DAWSON 1898, Proc. London Math. Soc. **29**, 519.

EGGENBERGER 1924, Mitt. d. Ver. schweiz. Versich.-Mathematiker Bern, **31**.
EULER 1938, Z. Phys. **110**, 450.
— und HEISENBERG 1938, Erg. exakt. Naturwiss. **17**, 1.

FELLER 1937, Math. Ann. **113**, 113.
— 1939, Acta Biotheoretica Leiden, **II**.
FRÉCHET 1938, Traité du calcul des probabilités et de ses applications, **I 3**, Livre II. Paris.
FROMAN and STEARNS 1938, Rev. Mod. Phys. **10**, 133.
FURRY 1937, Phys. Rev. **52**, 569.

GREENWOOD and YULE 1920, Journ. Roy. Stat. Soc. **83**, 255.

HEITLER 1936, Quantum Theory of Radiation, Oxford.
— 1937, Proc. Roy. Soc. A **161**, 261.
— 1939, Report on Progress in Physics **V**, 361.

IWANENKO and SOKOLOW 1938, Phys. Rev. **53**, 910.

JAHNKE and EMDE 1933, Tables of Functions, 2. ed., Leipzig und Berlin.

KHINTCHINE 1933, Asymptotische Gesetze der Wahrscheinlichkeitsrechnung, Berlin.
KOLMOGOROFF 1931, Math. Ann. **104**, 415.
— 1933, Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin.

LANDAU and RUMER 1938, Proc. Roy. Soc. A **166**, 213.
LOVELL 1939, Proc. Roy. Soc. A **172**, 568.
LUNDBERG, F. 1909, Verh. 6. internat. Kongr. Vers.-Wiss., Wien. **1**, 877.
LUNDBERG, O. 1940, On Random Processes and their Application to Sickness and Accident Statistics, Uppsala.

MØLLER and ROSENFELD 1940, Kgl. Dansk. Vid. Selsk. Skr. **XVII** no. 8.

NERESON 1942, Phys. Rev. **61**, 111.
NEWBOLD 1927, Journ. Roy. Stat. Soc. **90**, 487.
NORDHEIM 1938, Phys. Rev. **53**, 694.
NORDSIECK, LAMB and UHLENBECK 1940, Physica **7**, 344.

PEANO 1888, Math. Ann. **32**, 450.
PEARSON 1922, Tables of the incomplete Γ-Function, London.
PÓLYA 1930, Ann. l'Inst. H. Poincaré **I**, 117.

RASCH 1930, Om Matrixregning og dens Anvendelse paa Differens- og Differentialligninger, København.
ROSSI and JÁNOSSY 1939 a. Private communication to the author. See also the short note of Blackett and Rossi in Rev. Mod. Phys. **11**, 279.
— — — 1939 b. Rev. Mod. Phys. **11**, 281.

SCHLESINGER 1908, Vorlesungen über lineare Differentialgleichungen, Leipzig. Berlin.
— 1922, Einführung in die Theorie der gewöhnlichen Differentialgleichungen auf funktionentheoretischer Grundlage, 3. Aufl. Berlin und Leipzig.
SCHWEGLER 1935, Z. Phys. **96**, 62.
SERBER 1938, Phys. Rev. **54**, 317.
SNYDER 1938, Phys. Rev. **53**, 960.

TITCHMARSH 1932, Theory of Functions, Oxford.
TOEPLITZ 1909, Pal. Rend. **28**, 88.
TRUMPY 1943, Kgl. Dansk. Vid. Selsk. Skr. **XX** no. 6.
— and BJORDAL 1942, Bergens Museums Årbok. Naturvitenskapelig rekke no. 5.

VOLTERRA 1887, Mem. Soc. Ital. Sc.(3) **6** no 8.
— 1888, Pal. Rend. **2**, 69.
— 1902, Mem. Soc. Ital. Sc.(3) **12**, 3.

INDEX.

[The number refer to the pages. For definitions cf. the pages in italics.]